AD-A075 815 JOHNS HOPKINS UNIV BALTIMORE MD DEPT OF PSYCHOLOGY F/G 17/2
COOPERATIVE PROBLEM SOLVING BY LIKE- AND MIXED-SEX TEAMS IN A T--ETC(U)
AUG 79 P R MICHAELIS N00014-75-C-0131
UNCLASSIFIED TR-9 NL

1 OF 1
AD A075815

END
DATE
FILMED
11-79
DDC

LEVEL II

THE JOHNS HOPKINS UNIVERSITY

DEPARTMENT OF PSYCHOLOGY

AD A075815

DDC
RECEIVED
OCT 26 1979
E

This document has been approved for public release and sale; its distribution is unlimited.

TECHNICAL REPORT — 9

COOPERATIVE PROBLEM SOLVING BY LIKE- AND MIXED-SEX TEAMS IN A TELETYPEWRITER MODE WITH UNLIMITED, SELF-LIMITED, INTRODUCED AND ANONYMOUS CONDITIONS

AUGUST, 1979

ONR CONTRACT NUMBER N00014-75-C-0131
WORK UNIT NUMBER NR 196-135

DDC FILE COPY

79 10 26 016

Baltimore, Maryland 21218

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report - 9	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 Cooperative Problem Solving by Like- and Mixed-sex Teams in a Teletypewriter Mode with Unlimited, Self-limited, Introduced and Anonymous Conditions		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Paul Roller Michaelis		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-75-C-0131
9. PERFORMING ORGANIZATION NAME AND ADDRESS Paul Roller Michaelis, Department of Psychology The Johns Hopkins University Baltimore, Maryland 21218		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 196-135
11. CONTROLLING OFFICE NAME AND ADDRESS Engineering Psychology Programs Code 455 Office of Naval Research	11	12. REPORT DATE August 1979 13. NUMBER OF PAGES 53
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 14 TR-9 12 64		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Communication	Sex differences
Language	Telecommunication
Problem-solving	Teletypewriter communication

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It is well documented that men and women differ significantly in several important psychological respects. The pattern of these differences suggests the possibility that when a natural language problem-solving computer is built, it might have to adjust its language and problem-solving strategy to match that of the sex of its user. That possibility was examined in the experiment reported here.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified 400 605

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract - continued

The subjects were 96 Johns Hopkins University undergraduates, 48 men and 48 women. They were assigned to 48 two-person teams. One team member, the source, was given a completely assembled abstract wooden model and was required to assist the other member, the seeker, who had to build an identical model from the separate parts. The two team members were in different rooms, and all communication between them was via teletypewriters.

Twelve of the teams were all male, twelve were all female, twelve had a male source and female seeker, and twelve had a female source and male seeker. Half of the teams of each type were penalized for each word they used; the other half were not. Finally, on half of all teams the two participants were introduced to each other beforehand; on the other half they were not introduced and knew nothing of each other save what they could deduce from their teletypewriter conversation. This variable is referred to as the anonymity variable.

Performance was assessed on four classes of dependent measure: (1) the time required to assemble the model, (2) various measures of verbal productivity, (3) experimenter tabulations of various types of subject-to-subject information requests, and (4) responses to a questionnaire completed by the subjects at the conclusion of their session.

The only significant differences attributable to sex were that, on the average, teams with a female source were faster, used fewer words, and generated fewer of certain types of information requests. There were no significant differences attributable either to the main effect of sex of seeker, or to the interaction of sex of seeker with sex of source. Further, there were no significant differences attributable either to the main effect of anonymity or to the interaction of anonymity with any other variable. Therefore, this study found no important differences between men and women that would need to be taken into account when building a natural language problem solving computer.

There were several highly significant differences between the teams who were unlimited and those who were required to limit their word usage. The most interesting of these is that the limited teams assembled the model more quickly than did the unlimited teams. There were, however, no significant differences attributable to the interaction of this word usage factor with any other variable.

Cooperative Problem Solving by Like- and Mixed-sex Teams in a Teletypewriter Mode with Unlimited, Self-limited, Introduced and Anonymous Conditions

Paul Roller Michaelis

Based on a dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

This Research was supported in part by Contract Number N00014-75C-0131 between the Office of Naval Research and The Johns Hopkins University. Alphonse Chapanis is the principal investigator.

Reproduction in whole or in part is permitted for any purpose of the United States government.

Approved for public release; distribution unlimited.

Department of Psychology

1979

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or special
A

Table of Contents

	Page
Abstract	v
Introduction	1
Review of the Literature	2
Telecommunication Research at Johns Hopkins	2
Sex Differences in Cognitive Styles and Abilities	7
Sex Differences in Personality	9
Sex Differences in Interpersonal Relationships	10
Summary	11
Purpose	12
Method	13
Subjects	13
Experimental Facilities	13
The laboratory	13
Telecommunication equipment	13
Auxiliary equipment	15
Problem-solving Task	15
Experimental Design	17
Independent Variables	17
Dependent Measures	20
Time to solution	20
Number of words	20
Messages	21
Mean message length	21
Communication rate	21
Requests for information	21
Questionnaire data	22
Procedure	24
Scheduling of subjects	24
Preparation for the subjects	26
The arrival of the subjects	26
The instructions	26
The instructions	26
The experimental session	27

PRECEDING PAGE BLANK

Page

Results and Discussion 28

The Univariate Analyses of Variance 29
The Multivariate Analysis of Variance 34
Sex of Source 34

Time to solution 34
Number of words 34
Partner's understanding 36
Discussion 36

Word Usage 37

Time to solution 37
Number of words 37
Messages 38
Message length 38
Communication rate 38
Requests for information 39
Discussion 39

Sex of Seeker by Sex of Source by Anonymity 41

Job Roles 42

Seekers' clarity 42
Seekers' brevity 42
Discussion 42

Intercorrelations of Dependent Measures 42

Mistakes Made 45

Difficulties Encountered 46

Conclusion 47

Summary 47

Future Research 48

References 48

Abstract

It is well documented that men and women differ significantly in several important psychological respects. The pattern of these differences suggests the possibility that when a natural language problem-solving computer is built, it might have to adjust its language and problem-solving strategy to match that of the sex of its user. That possibility was examined in the experiment reported here.

The subjects were 96 Johns Hopkins University undergraduates, 48 men and 48 women. They were assigned to 48 two-person teams. One team member, the source, was given a completely assembled abstract wooden model and was required to assist the other member, the seeker, who had to build an identical model from the separate parts. The two team members were in different rooms, and all communication between them was via teletypewriters.

Twelve of the teams were all male, twelve were all female, twelve had a male source and female seeker, and twelve had a female source and male seeker. Half of the teams of each type were penalized for each word they used; the other half were not. Finally, on half of all teams the two participants were introduced to each other beforehand; on the other half they were not introduced and knew nothing of each other save what they could deduce from their teletypewriter conversation. This variable is referred to as the anonymity variable.

Performance was assessed on four classes of dependent measure: (1) the time required to assemble the model, (2) various measures of verbal productivity, (3) experimenter tabulations of various types of subject-to-subject information requests, and (4) responses to a questionnaire completed by the subjects at the conclusion of their session.

The only significant differences attributable to sex were that, on the average, teams with a female source were faster, used fewer words, and generated fewer of certain types of information requests. There were no significant differences attributable either to the main effect of sex of seeker, or to the interaction of seeker with sex of source. Further, there were no significant differences attributable either to the main effect of anonymity or to the interaction of anonymity with any other variable. Therefore, this study found no important differences between men and women that would need to be taken into account when building a natural language problem-solving computer.

There were several highly significant differences between the teams who were unlimited and those who were required to limit their word usage. The most interesting of these is that the limited teams assembled the model more quickly than did the unlimited teams. There were, however, no significant differences attributable to the interaction of this word usage factor with any other variable.

Introduction

At dawn on the morning of October 21, 1805, the British fleet, commanded by Lord Horatio Nelson from his flagship H. M. S. Victory, sighted the combined fleets of France and Spain about ten miles distant. What happened next is described in a contemporary biography of Nelson (Churchill, 1808):

> About twenty minutes after seven the enemy, who were in very compact order, wore in succession, and stood on the larboard tack with their heads toward Cadiz. They kept a good deal of sail set, steering about two points from the wind, with their topsails shivering; and formed a double line, the ships on one being opposite the intervals of the other, so that they were not crowded, though to our ships on their beam very little vacancy appeared between them. Their admirals did not show their flags, but the Santissima Trinidada being distinguished by her four decks, Lord Nelson ordered the Victory to be steered for her bow...and, about half an hour before the enemy opened their fire, the signal was made by telegraph, "England expects every man will do his duty." (p. 87)

Thus began the battle of Trafalgar. Because Nelson's signal to his fleet is one of the most famous telecommunications in history, it is interesting to see what Nelson's signal officer, Lieutenant John Pasco, had to say about it (Warner, 1968):

> His Lordship came to me on the poop, and about a quarter to noon said: "I wish to say to the fleet, ENGLAND CONFIDES THAT EVERY MAN WILL DO HIS DUTY," and he added, "You must be quick, for I have one more to make, which is for close action." I replied: "If your Lordship will permit me to substitute expects for confides, the signal will soon be completed, because the word 'expect' is in the vocabulary, and 'confides' must be spelt." Nelson replied in haste, and with seeming satisfaction: "That will do, Pasco, make it directly." (p. 180)

Although one might argue that Nelson's famous message was favorably altered by the device through which he communicated, the fact does remain that his original message was altered. We still have such problems today. Whenever we use devices to communicate, the devices themselves usually affect our communication in ways we do not yet completely understand. This is often no mere inconvenience; it is a genuine problem.

We now have the ability, through electronic devices, to communicate with people anywhere in the world. Harms (1974, chap. 3) gives a good description of the "telecomnet" that makes this possible. The tremendous technological achievement that the telecomnet represents may have made many of those associated with its development

overestimate its utility. For example, Licklider et al. (1968) predicted that, "In a few years, men will be able to communicate more effectively through a machine than face to face." Ochsman and Chapanis (1974) discuss this prediction, and state that it has not yet come to pass because of the "failure of scientists to attack systematically the many problems in this complex (man-machine communication) partnership."

Going beyond man-man communication via machines is the fascinating problem of actual man-machine dialogue. Will it ever be possible to build computers that would have all of the beneficial properties of Arthur C. Clarke's (1968) fictitious human-like HAL? While we undoubtedly already have the technological skills to build the hardware for a machine like HAL, as Chapanis (1971) pointed out, before HAL can become a reality we need a better understanding of how people themselves naturally communicate, and we need to know how their communication is affected by the devices through which they converse.

Horatio Nelson's problem is still with us, and only recently have we begun to understand its complex and subtle nature. In fact, it is not sufficient to study how people in general communicate, and to determine how their communication is affected by telecommunication devices. In order to gain a better understanding of the communication process, it is necessary to study different types or groups of people to look for group-attributable differences in the way they communicate. This means examining inter-group as well as intra-group dialogue, and seeing how the differences between groups are affected by telecommunication devices.

My dissertation experiment was a first attempt to examine the inter- and intra-group dialogue of these two fundamentally different groups of people: men and women.

Review of the Literature

This review will first provide a brief discussion of the literature on the human use of telecommunication systems. Following that, some of the research on the psychology of human sex differences will be examined. There is a tremendous amount of literature available on the psychology of sex differences, and this review cites only those references that are relevant to the study of sex differences as they affect the way people exchange factual information to solve problems.

Telecommunication Research at Johns Hopkins

Although the study of human communication itself is a major subspecialty within the field of psychology, much of what we now know about machine-mediated human interaction comes from research done at

Johns Hopkins under the direction of Professor Alphonse Chapanis. Most of the experiments at Hopkins have examined two-person teams who cooperatively exchanged factual information to solve realistic problems. Although the complete criteria for these problems appear elsewhere (e.g., Chapanis et al., 1972; Ochsman & Chapanis, 1974), for convenience they are summarized below:

1. The problems have sampled a wide range of behaviors.

2. They are representative of tasks for which telecommunications systems have been used, or could be used in the future.

3. The problems were not of the abstract or artificial type generally used to measure hypothetical psychological processes, but rather were the type people might encounter in everyday life.

4. They had definite, recognizable solutions that could be reached within an hour or less.

5. They required no special skills or specialized knowledge for their solution.

6. The problems were structured so that the two people on the team had to communicate in order to reach the solution. This was accomplished by giving the two team members complementary information folios. In general, one team member, the <u>seeker</u>, is presented with a problem he or she cannot solve without the help of the <u>source</u>, who has been given all of the information needed to solve the problem. The job role distinctions of seeker and source are important, and will be used throughout this paper.

As an illustrative example, the "Equipment Assembly Problem" developed by Chapanis et al. (1972) was structured so that one team member (the seeker) had all of the separate parts of a piece of equipment while the other member (the source) had the assembly instructions. The two had to communicate because neither could solve the problem alone. Another example is the "Close Scheduling Problem" (Weeks et al., 1974) in which the seeker is given a list of classes he must schedule for himself so that there are no time conflicts, doing so with the help of the source, who has the team's only list of class meeting times.

In general it has been shown that natural interactive human communication is characterized by extreme unruliness, both grammatically (Chapanis et al., 1977) and in terms of the number and kind of different words used (Michaelis et al., 1977). However, people have tremendous adaptive capacity, as demonstrated by Kelly and Chapanis (1977), who limited the vocabulary their subjects were allowed to use,

and Ford et al. (1979), who limited the number of words their subjects could use. In both these studies the restricted teams worked as fast and as accurately as did their unrestricted counterparts. Further, although interruptions are a common feature in normal human interaction, Chapanis and Overbey (1974) have demonstrated that they are not nearly as important as one might suspect.

How do communication devices affect human interaction? Five different modes of communication have been tested, either alone or in various combinations: handwriting, teletypewriting, telephone, close-circuit television, and, as a control, face-to-face.

A surprising finding in comparing communication modes is that there is very little difference between face-to-face communication and voice only (i.e., telephone). Apparently, seeing the person you are conversing with adds little to your ability to exchange factual information (Chapanis, 1973).

Interesting and highly significant differences appear when contrasting oral modes of communication (voice and face-to-face) with "hard-copy" modes (handwriting and teletypewriting). People communicating in oral modes are much faster at solving their problems, yet they use many more words (or tokens), more unique words (or types), more sentences, more messages, and have lower type-token ratios than people doing identical tasks in a hard-copy mode (Chapanis et al., 1972; Chapanis et al., 1977). In a recent experiment, Ford et al. (1979) demonstrated that these differences can be attributed to the fact that it is easier to talk than to write or type, and it is the basic inconvenience of handwriting or typing that provides the incentive to be concise. As an interesting sidelight to this, Weeks et al. (1974) compared the performance of skilled versus unskilled typists in a teletypewriter mode and found no significant differences in their performance, apparently because a relatively small percentage of one's time in these experiments is actually devoted to sending messages, and also because most people compose messages as they send them, thereby nullifying the advantage of being able to type copy rapidly.

There have also been studies examining the performance of various types or classes of subject. Parrish (1974) found that teams composed of college students were faster than teams composed of one college student and one high school student, but these mixed teams were faster than teams of high school students. Regarding the mixed teams, those with a college student as the source were faster than those with a high school student in this role. This suggests that the abilities of the source have a greater impact on the team's performance than do the abilities of the seeker.

Parrish (1978) has since conducted a small study in which he used a similar paradigm to compare incoming college freshmen to seniors at the University of California, Santa Barbara. Solution times did not

differ significantly between the two groups. This suggests that the differences he found in his earlier study are not attributable to educational level or age. However, the differences might have been due to intellectual ability. Desiring to examine intellectual ability more closely, Hoecker (1979) tested high school students to determine the effects of both verbal and spatial ability on their over-all ability to use telecommunications in problem solving. In general, subjects who scored high on the Differential Aptitudes Test subtest of spatial ability did better in his study than those who scored low. However, verbal ability, as determined by the Iowa Test of Basic Skills Reading Comprehension Test, could only be shown to affect performance on a problem that required the subjects to obtain their information from a written list.

One very basic human difference that has not been closely examined by the Hopkins group is sex. All but one of the Hopkins experiments have used subjects from just one sex, usually males. This was done partly out of convenience, as Hopkins has a predominantly male undergraduate population, but primarily because we have no real knowledge about the possible effects of the sex of the subjects on their behavior in a factually oriented communication situation. If there are effects on communication attributable to sex differences, simple random assignment to experimental conditions of subjects drawn from the general population would have increased the error variance in statistical analyses.

The experiment by Ford et al. (1979) was the first at Hopkins to include the variable of sex. Although the major focus of their study was on self-restricted word usage, half of their teams were composed of male subjects, and the other half of female subjects. The data contained no statistically significant main effects attributable to sex and only as many significant interactions of sex with another variable as would have been expected by chance. However, the primary focus of this study was not on differences between the sexes and, as a result, several independent and dependent variables relating to sex differences were not included in the experimental design. For example, there were no mixed-sex teams. Also, there were no clearly defined job roles (i.e., information source and seeker) but, rather, both subjects in the dyad started their session with roughly equal amounts of information. For these reasons, no "within teams" measures were taken and, as a result, any possible interaction of sex with job role could not be examined.

At the University of California, Santa Barbara, Parrish and Michaelis (submitted for publication) conducted an experiment to explore the behavior of like- and mixed-sex teams during cooperative problem solving. The performance of the teams was assessed on time required to solve the problem and on behavioral measures of activity such as the type described by Chapanis et al. (1972). The behavioral data were obtained through direct observation of the subjects while

they worked; every five seconds an observer would note which one of the 14 different previously defined behaviors the subject was engaged in. There were eight different single activities: Sending Only, Sending Pause, Receiving Only, Searching Only, Handling Parts, Making Notes, Waiting and Other; and six dual categories: Sending and Searching, Sending and Handling Parts, Receiving and Handling Parts, Receiving and Making Notes, Searching and Handling Parts, and Searching and Making Notes. Their sampling method allowed Parrish and Michaelis to estimate the actual amount of time the subjects spent in each behavior category.

Of the 14 behavioral categories sampled, six categories accounted for about 85 percent of the teams' activity, on the average. These six categories were: Sending Only, Receiving Only, Searching Only, Handling Parts, Making Notes, and Waiting.

As was the case with Ford et al. (1979), Parrish and Michaelis found no differences between their all-male and all-female teams. However, their mixed-sex teams, on the average, took over 50 percent longer to solve their problems than did teams of like sex. Because of the small number of subjects tested and the large amount of variability in performance, however, this difference was barely significant at $p < 0.05$.

Parrish and Michaelis tested two different communication problems. When Parrish and Michaelis designed these two problems, they had expected that males would be superior on one of them with females superior on the other. As it turned out, in the analysis of variance for time to solution, the interaction of sex with problem actually yielded an F ratio of less than one, so that there was not even a suggestion of significance. Further, an analysis of the behavioral data among groups of different sexual composition showed that, on the average, the different groups generally spent about the same percentage of time in each activity.

While the main effect of job role is not of great importance in this type of study, the interaction of job role with sex gives an indication of whether or not males and females do the identical jobs differently, and also whether or not their performance changes depending on the sex of their partner. This, in fact, is apparently what happened.

The most striking interaction was for the category of Searching Only. Male seekers, no matter with whom they were paired, averaged about 47 percent of their time in this category, and female seekers on like-sex teams averaged about 42 percent, while female seekers who had been paired with a male source averaged only 29 percent of their time Searching Only. However, these female mixed-team seekers spent more time in the dual categories of Sending and Searching and Searching and Making Notes than did the other seekers, causing the interactions for these two measures to reach significance also. These two activities

combined account for about 8 percent of the female mixed-sex seekers' time, as compared with about 1 to 2 percent for the other seekers. Parrish and Michaelis offer no hypothesis to explain these findings, but they do point out that these female-seeker mixed-sex teams averaged greater times to solution than did teams in the other groups.

Another finding of interest is that female sources averaged a greater proportion of their time in dual activities than did male sources, and this was true regardless of the sex of the seeker. Again, Parrish and Michaelis declined to offer an interpretation.

The major reason Parrish and Michaelis frequently did not interpret their findings is because their experiment was far too limited to allow them to do so. Among these limitations were:

1. The number of subjects was quite small, as only four all-male, four all-female and eight mixed teams were tested, with each team tested only once. As such, many possibly informative independent variables could not be included in the design.

2. The subjects were unpaid volunteers. Unlike the Hopkins experiments, this study at Santa Barbara provided no monetary incentive for the subjects to be either fast or accurate.

3. The subjects were in adjoining rooms, and communicated by passing handwritten notes to each other through a slot in the wall that separated them. The subjects were able to hear each other quite easily, and although they did not formally speak to each other during the course of the experiment, undoubtedly such things as laughter or groaning could have biased the results, especially with the mixed-sex teams.

4. The subjects seemed to take their tasks much less seriously than do Hopkins subjects. This may have been because of the relatively impoverished experimental conditions, but also because the Santa Barbara students, in general, seem to have less achievement motivation than Hopkins students. (This final observation is, of course, purely subjective.)

With these caveats in mind, it is apparent that Parrish and Michaelis left unanswered many questions about the communicative behavior of like- and mixed-sex teams during cooperative problem solving.

Sex Differences in Cognitive Styles and Abilities

There is a great deal of literature available concerning the psychological and sociological aspects of human sex differences. Some of that work has relevance to the study of sex differences in interactive problem solving.

It is well documented that if sufficiently large numbers of men and women are examined, differences can be demonstrated between them in their cognitive styles and abilities. In tests of verbal ability on subjects over the age of 16, when sex differences have been found, the differences have generally been in favor of females (see, for example, Monday et al., 1966-67; Rosenberg & Sutton-Smith, 1964, 1969). Typical of such findings are those by Droege (1967) who tested over 6,000 17-year-olds on the GATB (Verbal), and reported that the mean score for the females was 0.21 standard deviations higher than that of the males.

On the other hand, tests of quantitative and spatial ability for subjects over the age of 16 often yield higher scores for males (see, for example, Bieri et al., 1958; Rosenberg & Sutton-Smith, 1966, 1969; Stewart, 1976; Very, 1967). Illustrative data are those of Monday et al. (1966-67) who examined ACT math subscores of 237 18-year-olds, and found that the mean score of males was 0.47 standard deviations above that of the females. Further, it has long been known that "mathematical ability" is not a single factor by itself, but that it is, rather, made up of many independent components. Factor analyses by Blackwell (1940) and Very (1967), as well as others, show a different factor structure for males and females. For example, factor analyses done on males typically show the presence of one or more factors for spatial ability, while analyses done on women generally do not show these factors. As an interesting sidelight to this, the Harvard Project Physics (Walberg, 1969) administered physics achievement tests to both male and female high school physics students, and found that the males did better than the females on test items calling for visual-spatial skills, while females did better on verbal test items.

Because men and women differ in their cognitive styles and abilities, Maccoby and Jacklin (1974, p. 91) therefore assert that they also differ in the way they attack problems. It is certainly not impossible that some sort of mismatch in problem-solving strategies was involved when Parrish and Michaelis's mixed-sex teams took 50 percent longer than like-sex teams to solve identical problems.

In addition to looking at differences in mean scores between the sexes, psychologists have also been interested in differences in variability. Over a half-century ago, Terman (1925) reported that more boys than girls had I.Q.s in excess of 140, even though the mean scores for the two groups were about the same. This, on the surface, suggests that the variability in scores is higher for males than for females; however, Terman's subjects were not randomly selected, thereby making interpretation of this finding difficult.

Regarding tested verbal ability, more recently, Droege (1967) reported that the scores for 14- to 17-year-old females on the GATB (Verbal) were slightly more variable than those of males, although Gates (1961) and Svensson (1971) both reported 13-year-old males to be

more variable on verbal test items than females of the same age. Apparently, as Maccoby and Jacklin (1974, p. 120) point out, despite the generally higher verbal scores for females, the variability of males and females on these tests is about the same.

With regard to mathematical ability, however, there is strong evidence to suggest that males are more variable than females. Flanagan et al. (1961), Monday et al. (1966-67), and Svensson (1971) all report that males not only have higher mean scores than females, but also have greater standard deviations. This same finding is also reported by Keating (1974), but Keating's subjects were 396 junior high school students who had entered a science and math contest, either after having been selected by their teachers or through volunteering. Presumably, there was no sexual bias in the selection of these subjects, and even in this carefully selected group the males were still more variable and had a higher mean score than the females.

To summarize: On tests of verbal ability, college-age females tend to average scores slightly higher than those of college-age males, although the two groups exhibit roughly equal variability. On tests of mathematical ability, the males tend to do better than do the females and also, as a group, tend to be more variable. Factor analyses show that the two sexes tend to use different cognitive skills for solving problems.

Sex Differences in Personality

Not only are there sex differences in cognitive style and ability, but there are also sex-linked personality differences. Clearly, there are many such differences between males and females, so this review will examine only those that appear most likely to affect behavior in interactive problem solving.

The aspect of personality that may have the most bearing on performance of problem-solving tasks is the basic one of interest. It is not at all unreasonable to assume that people will do better at tasks they find interesting than at those they find dull. We must therefore ask this question: Does the subject in our experiment normally find tasks such as the one assigned to be of interest? Notice that the question did not ask if the subject found the actual experimental task interesting: nearly all subjects in previous Hopkins communications experiments, males and females alike, found participation in the experiments both interesting and enjoyable. Instead, the question seeks to find out if they are normally interested in this type of task because, if so, then they are more likely to have had prior experience with tasks that required similar cognitive abilities. As such, they might well have some advantage in the performance of the experimental task, or, at the very least, we might expect them to use a different strategy than someone with little prior experience.

As would be expected, males and females have, on the average, different interests. Monday et al. (1966-67) report that males tend to have more interest in technologically oriented fields, such as engineering, while females tend to have a greater interest in social, religious, and educational fields.

Holland's (1965) Vocational Preference Inventory has been used by Astin (1974) and Fox et al. (1976) to examine the vocational interests of youngsters who had been identified as being mathematically and scientifically precocious. The VPI has six categories of vocational preferences: Realistic, Investigative, Enterprising, Artistic, Social, and Conventional. The males, in general, tended to opt for vocations that were classed as Investigative or Realistic, while the females opted for Social or Artistic careers. One might reasonably expect that people who are "realistic" and "investigative" might be more naturally oriented toward problem-solving tasks.

As it turns out, in actual tasks involving problem solving, males, in general, seem more confident about their abilities than females (Carey, 1958; Crandall, 1969; Erlick & Le Bold, 1975; Feather, 1968; Rychlak & Lerner, 1965) and claim to be more sure of themselves in new situations (Schwartz & Clausen, 1970). Further, Deaux (1976) has reported that males are more likely than females to attribute success at a task to skill; the females tend to attribute their success to luck.

In general, we have seen so far that males and females differ not only in cognitive areas, but also in their interests and attitudes toward problem solving.

Sex Differences in Interpersonal Relationships

Cognitive patterns, interests, and levels of self-confidence are all intrapersonal attributes that could affect problem-solving ability. In addition, there are sex-linked interpersonal attributes that could affect performance in a situation requiring two people to work together to solve a problem.

One of the most widely recognized of these sex-linked attributes is that of aggression. Studies using questionnaires and standardized tests have found that males tend to be more aggressive and less patient than females (Barclay, 1970; Youssef, 1968). This finding is confirmed by experimental studies (Doob & Gross, 1968; Shomer et al., 1966; Youssef, 1968). In light of these findings, it is not surprising that females tend to "take orders" better than males (Maccoby & Jacklin, 1974, p. 371).

Most of these studies, however, were done before "feminism" became popular, so we need to ask if the female of today would still be less aggressive than males. Despite the studies suggesting that

sex-linked differences in personality may be attributed to differences in upbringing, there is strong evidence that differences in aggressive behavior are largely biological. It has been frequently demonstrated that the level of aggression, in both humans and primates, is directly correlated with the level of androgens in the blood, and that experimentally varying the level of androgens in an individual has a direct effect on his or her aggressive behavior (Barfield, 1976; Ehrhardt & Baker, 1973; Money & Ehrhardt, 1972). As such, we may expect that the previously noted sex differences in aggressive behavior are still valid.

The behavior of like- and mixed-sex dyads has also been studied, but this has generally been through the use of tasks and situations of little apparent relevance to studies that required people to cooperate in exchanging factual information. As an example, McNeal, McClintock and Nuttin (1972) examined the behavior of like- and mixed-sex pairs in the Prisoner's Dilemma game, and found that teams of mixed sex were less competitive than those of like sex. Bedell and Sistrunk (1973) also studied the Prisoner's Dilemma, and reported that the male-male and male-female pairs both cooperated more often than did the female-female pairs. Rapoport and Chammah (1965) reported that all-male pairs were more cooperative than mixed-sex or all-female pairs, while Tedeschi, Lesnick and Gahagan (1968) reported females to be more cooperative than males. Although the four reports presented above disagree in their findings, an examination of the bulk of the studies done on this topic suggests that males are, in general, more cooperative than females in this type of situation. It should be noted in this regard that some researchers feel that males are only physically more agressive, while females are aggressive in more subtle ways (e.g., "cattiness") that could affect how well and how easily they are able to cooperate with others (see, for example, Feshbach, 1970; McIntyre, 1972). The findings on this matter, however, are far from conclusive.

On the other hand, a series of studies reported by Rosenthal (1976) does bear directly on the question of sex effects in dyads when one person is giving instructions to another. Rosenthal's primary interest is in experimenter-subject interactions, but his finding could easily extend to the source-seeker interactions of our communication experiments. He reported that experimental sessions generally last longer when the experimenter and subject are of the opposite sex, as compared to when they are of the same sex. This was despite the fact that the experimenters, in all cases, were reading from identical instruction sheets, and were therefore giving verbally identical instructions. Further, in a personal communication, Rosenthal reported that even in the portion of the sessions where the experimenter was simply reading the instructions to the subject, mixed-sex experimenter-subject dyads more often than not took longer than like-sex dyads.

<u>Summary</u>

With regard to the study of telecommunication systems at Johns Hopkins, it has been demonstrated that when the devices are substituted

for face-to-face communication, these devices can affect the way that people "naturally" behave.

With regard to the study of human sex differences, it has been demonstrated that men and women differ in their cognitive styles, abilities, and strategies, in their likes, dislikes, and attitudes toward problem-solving situations, and in the way they interact with other people.

In his study of the ways in which an experimenter can inadvertently bias the results of an experiment, Rosenthal found that subjects generally worked faster if they had received their instructions from an experimenter of the same sex, even though, in all cases, the experimenters had been reading to the subjects from identical instruction sheets. Apparently, the people who had participated in Rosenthal's study behaved differently depending on the sex of the person with whom they had been paired.

In a study by Parrish and Michaelis which examined the behavior of two-person teams during cooperative problem solving, it was found that, on the average, mixed-sex teams took over 50% longer than like-sex teams to solve identical problems. This may simply be just one more example of subjects behaving differently because of the sex of their partner. However, the possibility exists that this finding could also be due, at least in part, to some sort of mismatch between men and women in the way they communicate to solve problems.

There is certainly ample evidence in the literature on sex differences to support the "mismatch hypothesis." If a mismatch of this type does exist, then it will not only have theoretical implications, but a practical one as well: the envisioned natural language computers of the future, if they are to be completely human-factored, may have to be strictly neuter, or perhaps even adjust themselves to the sex of their human user.

On the other hand, the subjects tested by Rosenthal and by Parrish and Michaelis were normal college-age men and women. Perhaps the mixed-sex pairs did not do as well because they had more on their minds than the assigned experimental tasks. While there is little or no evidence of social interaction in the verbal communication of these subjects, the experimenters made no attempt to control the possibly confounding effects of non-verbal communication. In fact, Rosenthal attributes his findings to non-verbal communication between his subjects, and this might well account for the findings of Parrish and Michaelis.

Purpose

The literature review cited many studies in which males and females were demonstrated to be significantly different on several psychological criteria. The purpose of my dissertation is to find out

whether males and females also differ in any practical way in how they exchange information over a telecommunication system. In particular, it seeks to discover differences that should be accounted for in the design of a natural language computer.

In order to accomplish this goal, my dissertation experiment was designed to answer the following specific questions: When two people of either sex solve a problem by communicating over a teletypewriter system, what differences, if any, would there be between the way the males and females communicated? Further, what differences would there be between the like- and mixed-sex pairs of communicators? What effect would it have, both in general and on the above differences, if these people communicated without knowing each other's sex? And, finally, what would happen to the communication if the communicators are given an incentive to be as concise as possible? The answers to these questions should provide an indication of whether there are any practical differences in the way men and women communicate to solve problems.

Method

Subjects

The subjects were 96 undergraduate students from the Johns Hopkins University, 48 males and 48 females. They were recruited through posters advertising an "interesting telecommunications experiment."

Experimental Facilities

The experimental facilities consisted of three major and fairly distinct items: the laboratory, the telecommunication system, and the auxiliary equipment.

The laboratory. This experiment was conducted in the telecommunication laboratory at Johns Hopkins. A diagram of the laboratory is shown in Figure 1. This experiment, like many others done here, assigned subjects to the job roles of source and seeker. For all sessions, sources were assigned to room A and seekers to room C; rooms B and D were not used.

Each subject's room contained a chair, a worktable, and equipment for several different telecommunications media, although only the teletypewriters were used in this study. I stayed in the observation room during the experimental sessions, and was able to observe the subjects through one-way mirrors.

Telecommunication equipment. The subjects communicated through two interconnected input-output writers (IBM Model #735). These input-output writers closely resemble regular electric typewriters, and have

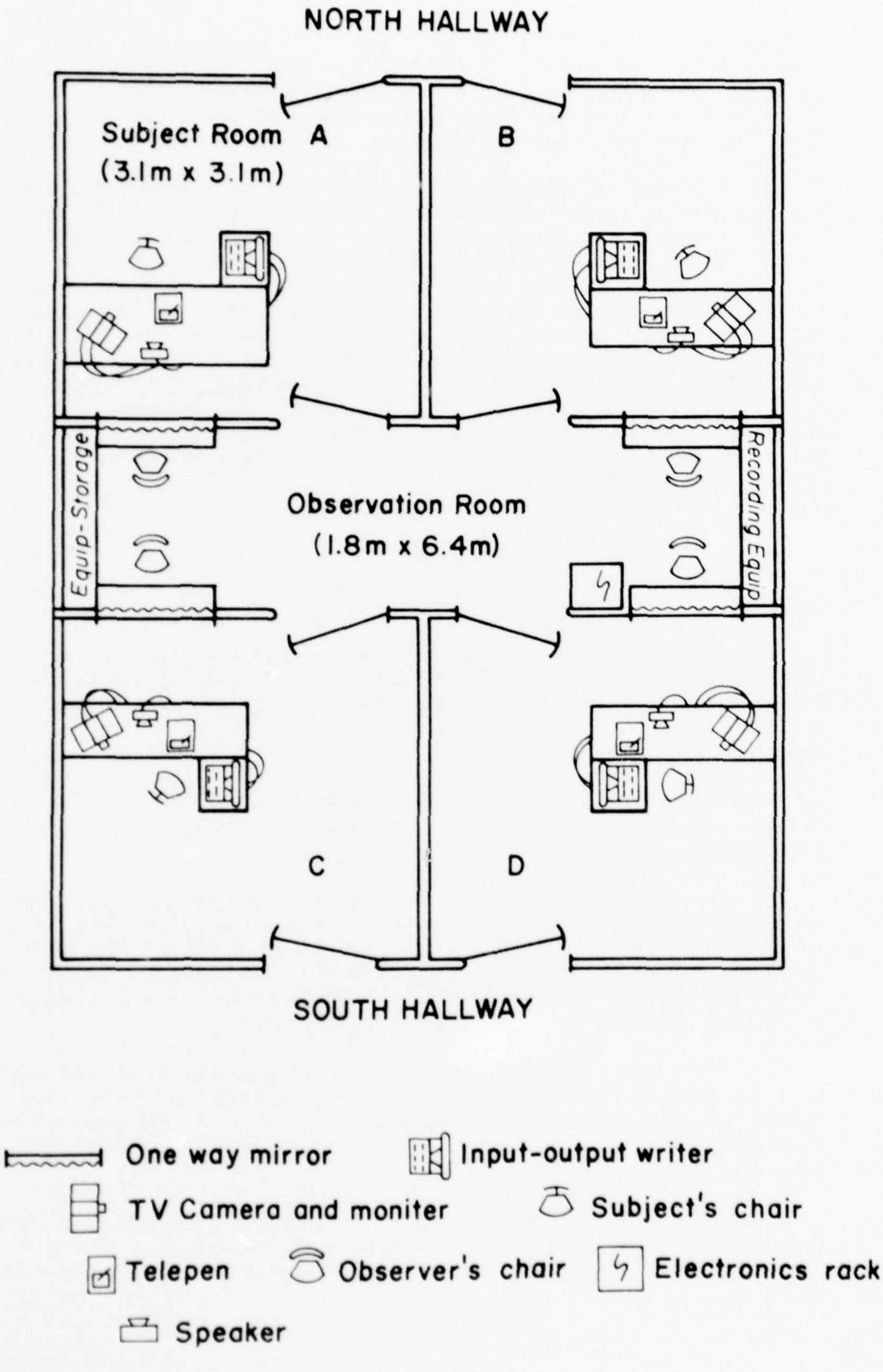

Figure 1. The telecommunication laboratory at Johns Hopkins.

standard typewriter keyboards, except for one extra key which is labeled "SEND."

This teletypewriter system was restricted to half-duplex operation. This means that only one subject could transmit at any given time. When a subject wanted to send a message, he had to first obtain control of the system by pressing the send key of his input-output writer. This freed the keyboard of his writer and put it in the "transmit" mode, while simultaneously locking the keyboard of his partner's writer by putting it in the "receive" mode. Whatever a subject typed appeared on his own terminal in red and on his partner's in black, thereby identifying the author of each message. Control of the system was confirmed for the subjects by a green pilot light on a small metal box mounted just above the input-output writer. A subject could keep control of the system until his partner pressed the send key on the other input-output writer. This feature allowed subjects to interrupt each other at any time.

Auxiliary equipment. The subjects received their instructions from a prerecorded tape played over a Sony Model TC-353 stereo tape recorder which was equipped with two detachable speakers. The tape recorder itself was kept in the observation room, and one speaker was placed in each of the two subject rooms.

An SGL Waber Model 24 switchable power output box was mounted near the input-output writer in room C, the seeker's room. The power for both input-output writers was routed through this box, as was the power for a 60-minute electric timer (Standard Electric Time Corporation Model S-60-ER). A discussion of how this system worked, and the rationale behind it, is provided in the Procedure section.

Problem-solving Task

The source was given a completely assembled Tinkertoy® model, and the seeker was given the parts from which he was required to build an identical model. A photograph of the model is shown in Figure 2.

The model was made of ten wooden dowel rods, color coded according to length, and eight wooden wheel-shaped connectors. All rods were one quarter inch in diameter; there were three green rods (7.5" in length), three blue rods (3.25" in length), and four yellow rods (2.125" in length). The connecting pieces had a diameter of 1.25" and a depth of .625". Six "SP" type connectors were used, these having one quarter-inch hole through their centers and eight quarter-inch holes around their circumferences. The rods fit snugly into these holes. Two "W" type connectors were also used, these having a three-eighths (rather than one-quarter) inch hole through their centers, four quarter-inch holes around the center hole, and four quarter-inch holes around their circumferences.

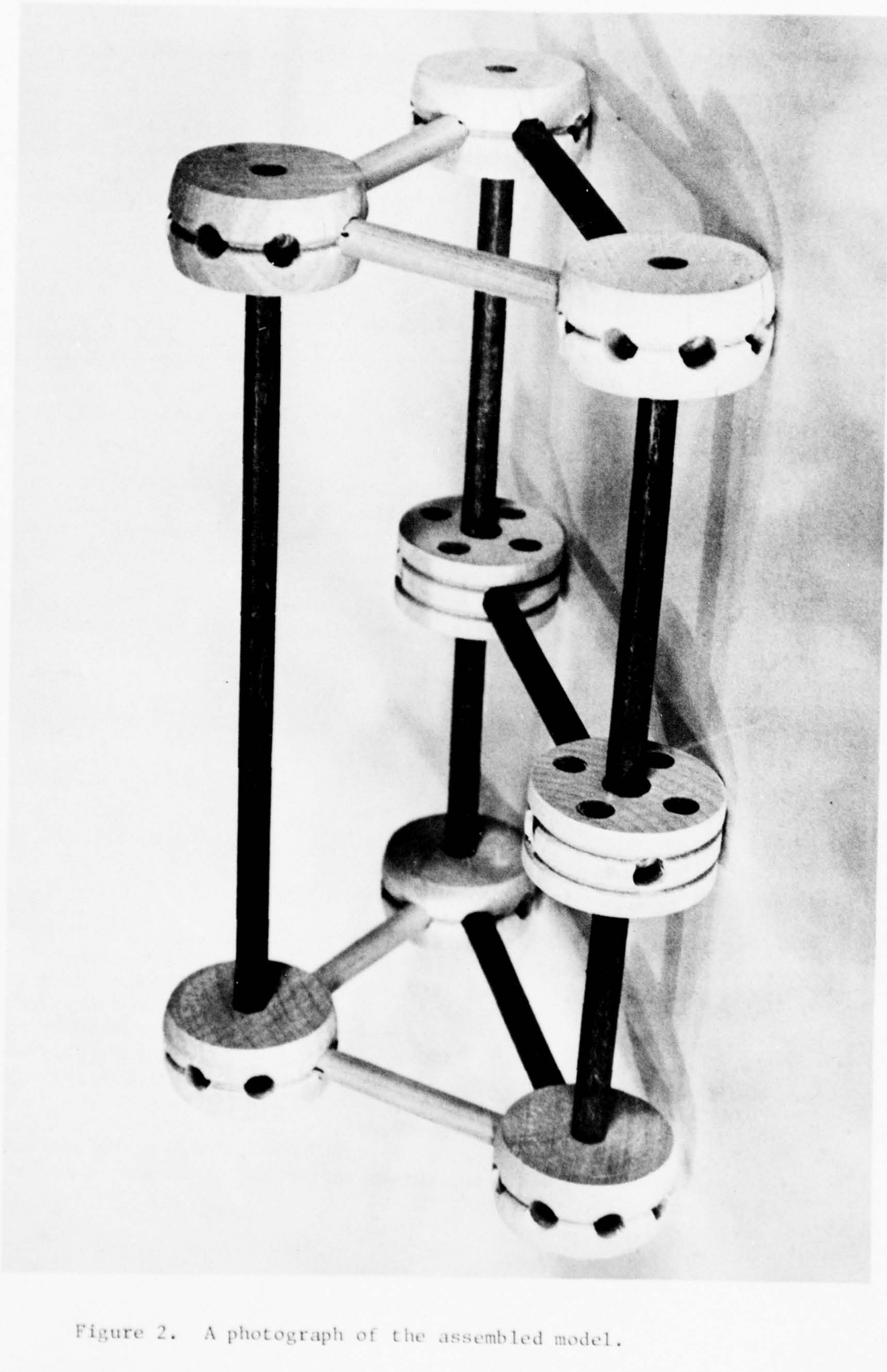

Figure 2. A photograph of the assembled model.

Experimental Design

In all, 48 two-person teams were tested in each of 16 different factor-level combinations (Table 1). There was one within-teams factor: job roles; and four between-teams factors: sex of seeker, sex of source, unlimited versus self-limited word usage, and introduced versus anonymous team members. Each factor had two levels.

There were two forms of the analysis of variance (Table 2). Form 1 was used for analyzing measures of individual performance (e.g., number of words), and Form 2 was used for measures of team performance (e.g., time to solution).

Independent Variables

The measure of job roles compares the performance of the two team members, the source and seeker. As has already been stated, the source was given a completely assembled model, and was required to assist the seeker, who had been given the separate parts from which he or she was required to build an identical model.

The factors of sex of seeker and sex of source are self-explanatory. Because these factors were crossed in the design, there were 12 all-male teams, 12 all-female, 12 mixed-sex with a male seeker, and 12 mixed-sex with a female seeker.

The factor of unlimited versus self-limited word usage, henceforth referred to simply as the word usage factor, was quite similar to the one used by Ford et al. (1979). Exactly what constitutes a word is explained in the Dependent Measures section, but, in general, I counted as a word any string of one or more typewriter characters separated from other strings by at least one blank space. The word usage factor involved allowing half of the teams to use as many words as they liked in solving the problem, while penalizing the other teams for each word they used. This was accomplished by the method I used for paying the subjects. All subjects received a base pay of $2.00 for participating in the experiment, plus a bonus for correctly solving the problem. The subjects on the unlimited teams received a fixed bonus of $1.50, while the subjects on the self-limited teams received a bonus of $3.25 minus one cent for each word that their team used. The average bonus awarded to self-limited subjects turned out to be slightly less than $2.00, and so was nearly the same as the $1.50 bonus received by the unlimited subjects.

Finally, there was the factor of introduced versus anonymous teams, henceforth referred to as the anonymity factor. For half of the teams in the experiment, the team members were allowed to see each other briefly at the beginning of the session. I introduced them by saying, "This will be your partner for the experiment," and then I separated them before they had a chance to speak to each other. The

Table 1

The Experimental Design

Problem Variables	Male Seeker: Male Source	Male Seeker: Female Source	Female Seeker: Male Source	Female Seeker: Female Source
Introduced				
Unlimited word usage	T_1, T_2, T_3			
Self-limited word usage				
Anonymous				
Unlimited word usage				
Self-limited word usage				T_{46}, T_{47}, T_{48}

At the University of California, Santa Barbara, Parrish and Michaelis (submitted for publication) conducted an experiment to explore the behavior of like- and mixed-sex teams during cooperative problem solving. The performance of the teams was assessed on time required to solve the problem and on behavioral measures of activity such as the type described by Chapanis et al. (1972). The behavioral data were obtained through direct observation of the subjects while

Table 2

The Two Forms of the Analysis of Variance, Showing Sources of Variation and Degrees of Freedom Associated With Each Source of Variation

Source of Variation	Form 1 d.f.	Form 2 d.f.
Sex of Seeker (SK)	1	1
Sex of Source (SO)	1	1
Introduced vs. Anonymous (A)	1	1
Unlimited vs. Self-limited Word Usage (W)	1	1
SK x SO	1	1
SK x A	1	1
SK x W	1	1
SO x A	1	1
SO x W	1	1
A x W	1	1
SK x SO x A	1	1
SK x SO x W	1	1
SK x A x W	1	1
SO x A x W	1	1
SK x SO x A x W	1	1
Teams/(SK-SO-A-W) combos.	32	32
Job Roles (JR)	1	-
JR x SK	1	-
JR x SO	1	-
JR x A	1	-
JR x W	1	-
JR x SK x SO	1	-
JR x SK x A	1	-
JR x SK x W	1	-
JR x SO x A	1	-
JR x SO x W	1	-
JR x A x W	1	-
JR x SK x SO x A	1	-
JR x SK x SO x W	1	-
JR x SK x A x W	1	-
JR x SO x A x W	1	-
JR x SK x SO x A x W	1	-
JR x Teams/(SK-SO-A-W) combos.	32	-
TOTAL	95	47

remaining teams were anonymous, and the team members did not meet each other until after the session was over. During the course of the session they knew nothing of each other save what they could deduce from each other's typewritten messages.

Dependent Measures

There were four classes of dependent measure: (1) the time required to solve the problem, (2) various measures of verbal productivity, (3) experimenter tabulations of various types of subject-to-subject information requests, and (4) responses to a questionnaire completed by the subjects at the conclusion of their session.

Time to solution. This is actually a measure of the total amount of time the communication system was open. When I turned on the system power at the start of the session, thereby allowing the subjects to communicate, an electric timer also received power. When the seeker was satisfied that the model was properly assembled, he turned the system off and thereby stopped the timer. Since this total time to solution is a measure of team performance, Form 2 of the analysis of variance (Table 2) was used for analyzing these data.

Number of words. While it may at first appear to be a simple task, it is actually quite difficult to determine how to count words, even in a teletypewriter mode. There is no agreement on this even among the Hopkins telecommunications group. The word counting rules used by Michaelis et al. (1977) are different from those used by Kelly and Chapanis (1977), and these are different from those used by Ford et al. (1978). In general, rules for word counting have been established and modified according to the context in which they will be used.

For this study, I usually counted as a word any string of one or more typewriter characters surrounded by spaces. Here are the specific rules I used:

1. In general, any string of letters, whether they spell anything or not, counted as a word. For example, "word," "wodr," and "owdr" would each count as one word. Single letters by themselves, such as "A," "B," or "C" would be counted as words.

2. Numbers count as words. For example, "3," "25," and "1/2" would each count as a word.

3. Special typewriter symbols, such as "@," "%," or "+" each count as individual words.

4. Punctuation marks not used solely for punctuation were counted as words. For example, "?????" and "......" each count as individual words.

5. Standard contractions, such as "can't" or "I'm" each count as one word.

6. Words run together or erroneously spelled were counted as their apparent constituent words. For example, "erron eously spaced" and "runtogether" would each count as two words.

7. Strikeovers count as words. For example, "~~word~~" would be counted as one word: "~~two words~~" as two.

The number of words used by each subject was counted independently by a research assistant and by me. Disagreements were rare, and were resolved to yield the values used in the analyses.

Messages. A message began when a subject obtained control of the communication system and started typing. The message ended when the subject's partner obtained control and typed something in return. It did not count as a message if a subject merely obtained control of the system and then lost it before typing anything. Messages had to be at least one typewriter character (i.e., one word) in length.

Although the source and seeker were taking turns sending messages, the number of messages sent by one does not have to equal the number sent by the other. If the subject who sent the first message is also the subject who sent the last, then he will have sent one more message than his partner. For this reason, Form 1 rather than Form 2 of the analysis of variance (Table 2) was used for this measure.

Mean message length. The mean message length is a derived measure, obtained by taking the total number of words used by a subject and dividing by the total number of messages generated by that subject.

Communication rate. Communication rate is another derived measure, obtained by taking the total number of words used by a subject and dividing by his team's time to solution. Notice that this is not a pure measure of communication rate, in that it is based on overall time to solution and not the amount of time actually spent communicating.

Requests for information. The five measures just described have the advantage of being quantitative. However, they only measure how much was said and how fast. They do not give us an idea of how well the subjects understood each other, a factor which gets at the essence of this study. I decided that the best way to judge this was to examine the record of each subject's communication with his partner to see how many and what kinds of questions were asked.

I should stress at this point that I was not interested solely in questions, at least according to the grammatical sense of that word. My interest lay in what I refer to as requests for information.

This is because the question, "Where does this piece go?" is, for my purposes, functionally the same as the simple imperative statement, "Tell me where this piece goes."

It proved to be extremely difficult to develop a reliable, unambiguous coding scheme to rate the various types of information requests. This was primarily due to the remarkably unstructured manner in which the subjects communicated. (This aspect of natural human communication is discussed in detail by Chapanis et al., 1977, and Michaelis et al., 1977.) After much trial and error, and with a great deal of help from my research assistant, the following coding scheme was developed:

1. Requests for New Information. (e.g., "What do I do next?")

2. Requests for Clarification of Previously Given Information. (e.g., "Could you go over that last part again?" or "Now I'm completely lost.")

3. Requests for Confirmation of Own Understanding or Status. (e.g., "Should those pieces form a triangle when I put them together?")

4. Requests for Confirmation of Partner's Understanding or Status. (e.g., "Do you understand?" or "Have you assembled the base of the model yet?")

5. Other. (e.g., "What time is it?")

The original typewritten protocols from all 48 experimental sessions were coded by my research assistant. I did not code them myself because of the possibility that my expectations could affect my judgment. However, to verify the reliability of the coding scheme, I independently coded five randomly chosen protocols, and came up with virtually the same tabulations as did my research assistant.

Questionnaire data. All subjects were asked to complete brief questionnaires at the end of their sessions. The questionnaire had been given to them inside a sealed envelope before the start of the session. The seeker was instructed to open the envelope and work on the questionnaire right after turning off the system; the source was instructed to do the same after hearing the power go off in his input-output writer.

The subjects on the introduced teams received the one-page questionnaire reproduced in Figure 3. For the subsequent analysis, the answers they gave to each of the five questions were converted to numerical values, one for the box on the far left to seven for the far right, for the seven possible answers on the rating scales.

QUESTIONNAIRE

This questionnaire is, of course, voluntary and strictly anonymous. DO NOT SIGN YOUR NAME. The five questions below simply ask you to rate your own and your partner's performance. Please check the appropriate box.

YOUR OWN MESSAGES:

1) Rate the CLARITY of your own messages to your partner:

Unclear and Obtuse ☐ ☐ ☐ ☐ ☐ ☐ ☐ Clear and Understandable

2) Rate the BREVITY of your own messages to your partner:

Unnecessarily Wordy ☐ ☐ ☐ ☐ ☐ ☐ ☐ Brief and Concise

YOUR PARTNER'S MESSAGES

3) Rate the CLARITY of your partner's messages to you:

Unclear and Obtuse ☐ ☐ ☐ ☐ ☐ ☐ ☐ Clear and Understandable

4) Rate the BREVITY of your partner's messages to you:

Unnecessarily Wordy ☐ ☐ ☐ ☐ ☐ ☐ ☐ Brief and Concise

....

5) Rate how well you and your partner were able to work together:

The two of us made a terrible team. ☐ ☐ ☐ ☐ ☐ ☐ ☐ The two of us made a great team.

Figure 3. The first page of the questionnaire.

The subjects on the anonymous teams received a two-page questionnaire. The first page was the same as the questionnaire given to the introduced teams (Figure 3). The second page is shown in Figure 4. I was interested in their answers to the second question on this page. This is because the major focus of my study is on sex effects, and, in interpreting my findings, I need to know if my anonymous teams really were anonymous with regard to sex. Note that this question is actually composed of three distinct parts: (1) What is your partner's sex? (2) How sure are you? and (3) What is the basis for your decision? This allows me not only to judge the effectiveness of my anonymity condition, but also provides me an insight into the way people form judgments of others when they are communicating over a teletypewriter system.

The introduction to this page of the questionnaire, as well as questions 1, 3, and 4, had as their primary purpose to prevent the subjects from realizing that mine was a study of sex differences. I did not analyze the answers to these questions, except as they related to the sex question.

Procedure

The subjects were solicited by posters, put up at various locations on the Johns Hopkins University campus. The posters directed interested students to a sign-up sheet which had been posted on the door to my office. The sheet asked them for their names and telephone numbers.

Scheduling of subjects. Every Sunday evening during the course of the experiment, I randomly selected fourteen to twenty potential subjects from the sign-up list and telephoned them. I described my experiment as involving two-person teams communicating via teletypewriters to solve a problem, asked them if they were interested and, if so, when during the upcoming week they would be available to participate. Once I had this information from all of the available subjects, I randomly assigned them to teams within the various conditions. Job roles in the mixed-sex teams were, of course, assigned on the basis of sex, and in the like-sex teams were assigned on the basis of a coin flip. After I had paired the subjects into teams, I examined the time schedules they had given me to find a mutually agreeable time for the two subjects. I then telephoned them to tell them when and where to report.

Subjects who were to be sources were told to report to Ames Hall room 254 (room A in Figure 1), while future seekers were told to report to room 216 (room C). The subjects were told that they would find the door to their room open, and that they were to come in, close the door, take a seat, and wait for me to come in and speak with them.

Notice in Figure 1 that the two subject rooms are on different hallways. This helped prevent the subjects from seeing each other in advance, and from realizing that they were to participate in the same experiment.

Some scientists have suggested that the teletypewriter mode of communication is too "impersonal," especially when compared with other telecommunication modes, such as telephone or closed-circuit television. Although teletypewriters may <u>seem</u> to be more impersonal than these other modes, this has yet to be scientifically demonstrated.

Both you and your partner are now working on this questionnaire, which is my way of trying to find out just how impersonal teletypewriters are. First, please answer these questions so we can compare your description of yourself with your partner's description of you:

Your year in school:________________ Your academic major:____________________

Your sex:______________ Your hometown:_____________________________________

Now, try to answer the questions below based solely on what you have learned about your partner during the course of your teletypewriter conversation. Please indicate how sure you are of each answer by checking the appropriate box in the rating scale; if your answer is more than "Just a guess" then please write on the back of this page what you based your decision on.

....

1) What is your partner's year in school: ____________________________
How sure are you?

Just a guess ☐ ☐ ☐ ☐ ☐ ☐ ☐ Absolutely certain

2) What is your partner's sex? ______________________
How sure are you?

Just a guess ☐ ☐ ☐ ☐ ☐ ☐ ☐ Absolutely certain

3) What is your partner's academic major? ____________________________
How sure are you?

Just a guess ☐ ☐ ☐ ☐ ☐ ☐ ☐ Absolutely certain

4) In what geographic region of the U.S. is your partner's hometown?

_________________________________ How sure are you?

Just a guess ☐ ☐ ☐ ☐ ☐ ☐ ☐ Absolutely certain

Figure 4. The second page of the questionnaire.

Preparation for the subjects. Before the subjects were to arrive, I put on the worktables in each room a standard subject consent form, a sheet of blank paper that the subjects could use for notes, and three sharp pencils. I left the doors to the hallways wide open, and the doors to the observation room shut. I positioned myself between the two one-way mirrors in the observation room and waited for the subject to arrive.

The arrival of the subjects. Even though the two subjects on a team were both assigned the same arrival time, out of all 48 teams, I did not have one where the subjects arrived within two minutes of each other.

If the team was to be introduced, when the first subject arrived I poked my head through the observation room door and said, "We'll be ready to begin as soon as your partner arrives." When the second subject arrived, I allowed him to briefly stand in the doorway to the first subject's room as I introduced them with: "This is your partner for the experiment." I then immediately escorted the second subject to his seat in the other room. Notice in Figure 1 that at this point, even with the observation room doors open, the subjects could not see each other.

I then stood between the two open observation room doors, said "I'll start playing you the tape with your instructions on it," and shut the doors.

Obviously, there was no introduction with the anonymous teams, although most other aspects of the procedure were the same. The difference is that when the second subject arrived, I simply poked my head through the door to his room and informed him that his partner had already arrived so I was going to begin playing the instruction tape. I closed the door after me, quickly informed the first subject that we were ready to begin, and then started playing the tape.

The instructions. In all previous Hopkins telecommunication experiments, the experimenter has read the instructions to the subjects face-to-face. However, keeping in mind Rosenthal's (1976) findings about the effects of unintentional experimenter-subject interaction, I elected to prerecord my instructions on tape to insure that all teams heard instructions that were identical, both verbally and nonverbally.

There were two sets of instructions, one for the unlimited teams and one for the self-limited. These differed only in their explanation of the bonus payment system.

The instructions began by thanking the subjects for participating in my experiment, the purpose of which, I said, was to see what happens when people solve problems by communicating over a teletypewriter system. I went on to explain the nature of the problem they had to solve (i.e., the assembly of the model), and what their job roles would be.

Next, I explained the payment system. I did not at this time fully explain my rules for word counting to the self-limited subjects. Rather, I told them that I would shortly be giving them a list of rules to which they could refer during the course of the session.

I then provided a thorough explanation of how to use the teletypewriter system.

After that, I explained to the subjects that it would ruin my analyses if, during the session, they engaged in any sort of "chit-chat," or especially if they asked for or gave any information not directly related to assembling the model. I told them that, if they wanted, they would have a chance to get to know each other or play with the teletypewriters after the session was over.

Finally, I concluded the instruction tape by telling them that I would be in to talk with them shortly to answer their questions and provide them with the material they would need for the experiment.

After the tape was over, I poked my head into the seeker's room and said, "I'll be with you in a few minutes. I want to talk to your partner first." In addition, if the team was to be self-limited, I gave the seeker a copy of word counting rules and suggested he examine them until I returned. This list of word counting rules was identical to the list of rules presented in the Dependent Measures section of this paper.

I then entered the source's room, carrying with me a box containing the assembled model and the envelope with the questionnaire. If the team was self-limited I also brought a copy of the word counting rules and had the source examine it before I did anything else. I allowed the source to ask questions, which I answered by paraphrasing what had already been said on the taped instructions. I put the box and the envelope on the worktable and said, "I want to go talk to your partner now. Please don't examine the model until you hear the system come on, and don't examine the questionnaire until the system goes off." I then asked the subject to sign the consent and payment forms and give them to me.

I followed essentially the same procedure with the seeker, except, of course, I gave the seeker a box containing separate parts rather than an assembled model. Also, before leaving, I made certain that the seeker knew when and how to shut the communication system off.

The experimental session. The actual experimental session began as soon as I turned on the system power. The motors in the input-output writers make a clearly audible noise when they first receive power, and this was the signal for the subjects to begin work. From my vantage point in the observation room, I could watch the subjects while they worked, and could even read what they were typing.

The timed portion of the session ended when the seeker shut off the system, at which point the subjects began work on their questionnaires. When they were done, I permitted them to meet each other and discussed various aspects of the experiment with them. I thanked them when they left, and asked them to encourage others to sign up for the experiment. The subjects subsequently received their payment checks through the campus mail, usually within a few days.

Results and Discussion

All 48 teams reached a solution. Only one team, an all-female self-limited anonymous team, failed to reach the correct solution. In other words, the seeker on that team failed to build an exact replica of the source's model. Because 47 of the 48 teams reached the correct solution, accuracy of solution is not a useful dependent measure for this study, and will be ignored in the remaining analyses. However, before any further analyses are done, it is necessary to examine the questionnaire data from the anonymous subjects to determine if they genuinely were anonymous.

As was explained in the Questionnaire data section, the anonymous team members were asked about the sex of their partner, how sure they were, and what they based their decision on.

There were 48 anonymous team members, 31 of whom correctly guessed the sex of their partner. Assuming that this is a simple binomial choice, with the likelihood of one answer equaling the likelihood of the other, the probability of getting by chance 31 or more correct answers equals 0.03. This is less than the traditional cut-off level for statistical significance of 0.05, but not by much.

A closer examination reveals that the correct responses were fairly evenly divided between job roles, 16 out of 24 correct for seekers, 15 out of 24 correct for sources. However, broken down by sex, there were 13 right out of 24 for males, as compared with 18 right out of 24 for the females. The probability of getting by chance 18 or more correct out of 24 equals 0.01, which is statistically significant.

In addition to deciding about their partner's sex, the subjects had to indicate how sure they were of their answer on a scale of one to seven, where one equals "just a guess" and seven equals "absolutely certain." I took the numerical rating from this scale and recorded it as a positive number if they had correctly guessed their partner's sex, and as a negative number if they had not. For example, if a subject incorrectly guessed his partner's sex, and checked the third box on the scale, I recorded a -3 for him. In this way I constructed a scale that runs from -7 to +7, with zero as the mid-point. On this scale the average score for the subjects was 0.23, with a standard deviation of 4.12. A t-test revealed that this mean score was not significantly

different from zero. Broken down by sex, the females averaged 0.96 on this scale, while the males averaged -0.5. Neither of these were significantly different from zero, nor were they significantly different from each other.

Out of the 48 subjects, 36 rated their assuredness as 2 or higher on the 7-point scale. In other words, 36 subjects felt that their judgment of their partner's sex was more than "just a guess." Of those 36, 15 were wrong. Once again making use of the binomial tests, the probability of getting by chance 15 or fewer wrong answers out of 36 is greater than 0.2, clearly not significant.

The 12 remaining subjects replied with what they considered to be just guesses. However, 10 of these 12 made the right guess. The probability of having, by chance, two or fewer wrong answers out of 12 is less than 0.02. So, the subjects who were more sure of their answers did far worse than those who said that they had simply guessed, which is why some of the simple binomial tests reached significance, although the t-tests on the corresponding rating scale scores did not.

The subjects who answered with "more than just a guess" were asked to explain why. The seeker on one all-male team replied that his partner "sounded educated," and was therefore female. A male on another team wrote, "My partner used the word 'botched' and women don't use that word." His partner proved to be the exception to that rule. Still others rated their partners solely on the basis of typing ability, with the popularly held stereotype apparently being that females are better typists than are males.

It is my belief that, on the average, the subjects who "just guessed" were basing their responses on indicators of sex that were far more subtle, yet apparently more accurate than those used by the other subjects. The findings also suggest that females may be better than males at detecting and properly interpreting these indicators. However, the evidence for these assertions is far from conclusive.

Further, and of more importance to this study, the 48 anonymous subjects averaged only 0.23 on the -7 to +7 judgment scale, confirming that, unlike their introduced counterparts, the anonymous subjects interacted without at any time knowing the sex of their partners.

The Univariate Analyses of Variance

After confirming the effectiveness of the anonymous condition, the next step was to perform univariate analyses of variance on each of the fifteen dependent measures used in this study. The results of the analyses done on the time and verbal measures are shown in Table 3, the different types of requests for information in Table 4, and the questionnaire data in Table 5.

Table 3

The p Values from the Analyses of Variance Performed on the Time and Verbal Productivity Measures

The values shown are for the upper end of the confidence interval: $0.05 > p > 0.025$; $0.025 > p > 0.01$; $0.01 > p > 0.005$; $0.005 > p > 0.001$.

	Dependent Measures				
Source of Variation	Time to Solution	Words	Messages	Message Length	Communication Rate
Sex of Seeker (SK)					
Sex of Source (SO)	.05	.025			
Anonymity (A)					
Word Usage (W)	.005	.001	.001	.001	.001
SK x SO					
SK x A					
SK x W					
SO x A					
SO x W					
A x W					
SK x SO x A					
SK x SO x W					
SK x A x W					
SO x A x W					
SK x SO x A x W					
Job Roles (JR)	----	.001		.001	.001
JR x SK	----				
JR x SO	----				
JR x A	----				
JR x W	----	.001	.001		.005
JR x SK x SO	----		.05		
JR x SK x A	----				
JR x SK x W	----				
JR x SO x A	----				
JR x SO x W	----				
JR x A x W	----				
JR x SK x SO x A	----				
JR x SK x SO x W	----				
JR x SK x A x W	----				
JR x SO x A x W	----				
JR x SK x SO x A x W	----				

Table 4

The p Values from the Analyses of Variance Performed on the Different Types of Information Requests

Source of Variation	New Information	Clarification	Own Understanding	Partner's Understanding	Other
	Dependent Measures				
Sex of Seeker (SK)				.05	
Sex of Source (SO)				.05	
Anonymity (A)					
Word Usage (W)	.001		.025	.001	
SK x SO					
SK x A					
SK x W				.025	
SO x A					
SO x W					
A x W					
SK x SO x A					
SK x SO x W					
SK x A x W				.025	
SO x A x W					
SK x SO x A x W					
Job Roles (JR)	.001	.001	.001	.001	
JR x SK					
JR x SO					
JR x A		.05			
JR x W	.001		.025	.001	
JR x SK x SO					
JR x SK x A					
JR x SK x W					
JR x SO x A					
JR x SO x W					
JR x A x W					
JR x SK x SO x A					
JR x SK x SO x W					
JR x SK x A x W				.05	
JR x SO x A x W					
JR x SK x SO x A x W					

Table 5

The p Values from the Analyses of Variance Performed on the Questionnaire Data

	Dependent Measures				
Source of Variation	Seekers' Clarity	Seekers' Brevity	Sources' Clarity	Sources' Brevity	Team Performance
Sex of Seeker (SK)	.025				
Sex of Source (SO)					
Anonymity (A)					
Word Usage (W)					
SK x SO					
SK x A					
SK x W					
SO x A					
SO x W					
A x W					
SK x SO x A				.05	
SK x SO x W					
SK x A x W					
SO x A x W					
SK x SO x A x W					
Job Roles (JR)	.001	.01			
JR x SK					
JR x SO					
JR x A					
JR x W	.05				
JR x SK x SO			.05		
JR x SK x A					
JR x SK x W					
JR x SO x A					
JR x SO x W					
JR x A x W					
JR x SK x SO x A					
JR x SK x SO x W					
JR x SK x A x W					
JR x SO x A x W					
JR x SK x SO x A x W					

As can be seen from these three tables, 449 separate F-ratios were computed. By definition, about 22 of the 449 F's would be expected to reach the 0.05 level of significance by chance. In actual fact, 36 reached significance, and it is not immediately clear which of these 36 indicate the presence of a genuine effect and which are spurious significances.

Nine of the significances are for the main effect of job roles. Eight of these are significant beyond the 0.001 level. However, this is neither surprising nor terribly interesting. All of the previous Hopkins studies that have had specific job roles have reported highly significant differences attributable to this factor. Although the exact nature of the differences in this experiment will soon be discussed in detail, in general they mean simply that people who build models do not behave the same way as those who give instructions.

There were eleven significant interactions of job roles with either between-teams main effects or between-teams effect interactions. In seven of these cases, the corresponding between-teams main effect or interaction was also significant by itself. As a general simplification for these seven instances, not only did the particular between-teams factor affect overall team performance, but it also affected the source and seeker differently, hence the significant interaction with the job roles factor.

The four remaining significant interactions with job roles barely reached significance ($0.05 > \underline{p} > 0.025$), and may therefore be chance effects. This is not saying the seven other significant job roles interactions are not chance effects. They could easily be. In order to demonstrate that they are probably not chance effects, it is necessary to demonstrate that the between-teams component of the interaction was itself not a chance effect.

All previous Hopkins telecommunications studies have used a relatively simple method for determining chance effects. The method had been to ask, "For this source of variation, what is the overall probability of obtaining by chance the number and magnitude of significant probabilities that were observed?" If the probability level thus computed was greater than a previously chosen Alpha level (usually 0.05) then all significant univariate probabilities for that source of variation were assumed to be chance effects.

This method has one important underlying assumption: The dependent measures must be uncorrelated. Unfortunately, this is not by any means a valid assumption, as will be demonstrated in the <u>Intercorrelations of Dependent Measures</u> section. Because the measures are intercorrelated, the probabilities of their reaching significance are linked, thereby invalidating the above described method of determining chance effects. A set of several dependent measures may show significances solely because all these measures are simply different ways of

describing the same single underlying factor. It follows that this single factor might cause a chance effect, and this would show up as significances for all the different dependent measures that describe it.

The Multivariate Analysis of Variance

Clearly, what is needed is an analysis of variance done on these underlying factors rather than on the actual dependent measures. Greatly simplified, this is what is done by the multivariate analysis of variance, or MANOVA, which computes a single F-ratio for each source of variation. This F is computed using a linear combination of all applicable dependent measures, with the factor weights in the linear combination being chosen so as to maximize the resulting F-ratio.

In general, if the multivariate F-ratio for a certain source of variation fails to reach significance, then all significant univariate F's for that source of variation are assumed to be chance effects. On the other hand, if a multivariate F for a source of variation does reach significance, then all corresponding significant univariate F's are regarded as being genuinely significant.

This is the method I used for determining which of the between-teams effects in Tables 3, 4, and 5 were genuine and which were chance. The results of the MANOVA are summarized in Table 6, which shows the multivariate F and corresponding p level for all between-teams sources of variation. As can be seen in Table 6, only three sources of variation were significant beyond the 0.05 level: sex of source, word usage, and the sex of seeker by sex of source by anonymity interaction.

Sex of Source

The sex of source was significant in three univariate analyses of variance: time to solution, number of words used, and number of requests for information on partner's understanding or status.

As will be seen in Table 8, in the Intercorrelations of Dependent Measures section, the mean within-cell intercorrelation of time to solution with words was moderately high ($r = 0.460$, $p < 0.01$). On the other hand, the number of requests for information about the partner was only marginally correlated with the number of words ($r = 0.341$, $p < 0.06$), and was uncorrelated with time to solution ($r = 0.009$). This suggests that at least two independent underlying factors accounted for the significant multivariate F for sex of source.

Time to solution. The average time to solution for teams with a male source was 26.7 minutes, while it was only 21.5 minutes for teams with a female source.

Number of words. Teams with a male source used an average of 383 words to solve their problem, while teams with a female source used an average of only 291 words to solve the identical problem.

Table 6

Summary of the MANOVA Performed on the Between-teams Sources of Variation

Sources of Variation	F-ratio	p less than
Sex of Seeker (SK)	0.771	0.691
Sex of Source (SO)	2.935	0.016
Anonymity (A)	1.180	0.364
Word Usage (W)	16.533	0.001
SK x SO	1.035	0.466
SK x A	1.472	0.216
SK x W	1.309	0.290
SO x A	1.684	0.145
SO x W	0.535	0.887
A x W	0.500	0.910
SK x SO x A	2.275	0.049
SK x SO x W	0.760	0.701
SK x A x W	0.862	0.610
SO x A x W	1.251	0.322
SK x SO x A x W	1.289	0.301

NOTE: All of the above F-ratios have 15 degrees of freedom in the numerator and 18 in the denominator.

Partner's understanding. On teams with a male source, the team members questioned each other's understanding an average of 4.4 times per team, while the teams with a female source averaged only 2.6 questions of this type per session.

Discussion. Although I had no firm expectations about what my study would find with regard to sex differences, I certainly did not expect the results to turn out as they did. The sex of the seeker made no apparent difference, and the sex of seeker by sex of source interaction failed even to approach significance. However, the main effect of the sex of the source did make a difference, but not the one that might have been expected: On a problem-solving task requiring spatial and mechanical skills, it was the females who were superior, although males supposedly are better at problem solving and have higher spatial and mechanical ability.

On the other hand, this was a telecommunications experiment, and females, on the average, do tend to have slightly higher verbal ability. However, this is unlikely to be the sole reason for the differences between the male and female sources in my study. Recall that in a previous experiment, verbal ability, as measured by a standardized test, could not be shown to affect performance on a problem-solving task similar to the one I used (Hoecker, 1979).

I can only offer, based on my observations, an educated guess as to why the female sources were superior in this experiment. Although I did not maintain records of this during the course of the experiment, I did notice that the seekers on male-source teams improperly assembled and then had to disassemble the model more often than did the seekers on female-source teams. That this took place is sometimes apparent from reading the transcripts of their conversations. On the average, the female sources tended to give instructions that were a bit longer, but apparently more thorough and less ambiguous than those given by the males.

The seekers on the male-source teams had a greater tendency to become confused by the instructions they had received. Often they would try to assemble the model according to their instructions, finally realize that it was not working, and then ask for help. (The seekers on male-source teams averaged 2.6 requests for clarification per session, as opposed to 1.9 for seekers on female-source teams; $0.1 > p > 0.05$.) The male sources on these teams would then go back over the different stages of the model's assembly, asking each time something like, "Did you understand that part?" Each of these, of course, counted as a request for information about the partner's understanding or status. This accounts for the significant difference between male- and female-source teams for this type of information request. The fact that the male sources had to explain some things over again accounts for their teams' longer times to solution and greater verbal output.

The underlying reasons for these results are not at all obvious. As was discussed earlier, males tend to approach problem-solving tasks with greater confidence than do females; perhaps in this case they were simply over-confident to the point of being careless. However, something more subtle than that may have been happening here. Given that, on the average, females tend to be poorer than males at mechanical and spatial tasks (and, perhaps more importantly, also believe themselves to be poorer), it is possible that the females in this study perceived the model to be more complicated than did the males, and therefore made more of an effort to describe clearly and in greater detail. This, perhaps in conjunction with the previously discussed differences in general confidence and verbal skills, may be the reason for the female sources' superiority in this study. However, this assessment represents pure speculation on my part, as I have no hard data to back me up. It will require further experimentation to fully explain the sex differences found in this study.

With regard to the absence of a significant sex of seeker by sex of source interaction, I believe that this can be attributed to the absence of nonverbal communication in teletypewriter dialogue.

Word Usage

The main effect of word usage was significant beyond the 0.05 level in eight of the univariate analyses of variance. These eight consisted of the analyses on all five of the time and verbal measures, plus analyses dealing with three of the five types of requests for information: requests for new information, to confirm own understanding or status, and to confirm partner's understanding or status.

Time to solution. The unlimited teams took an average of 27.6 minutes to solve their problem, while the self-limited teams took an average of 20.5 minutes to solve the identical problem.

Number of words. On the average, the unlimited teams used 545.5 words, as compared with the self-limited teams, who used 128.6 words. Also, averaged across all teams, seekers used an average of 103.8 words, while sources used 233.3, thereby accounting for the significant job roles main effects.

There was also a highly significant job roles by word usage interaction. Unlimited seekers used an average of 177.8 words while unlimited sources used an average of 367.7 words. On the other hand, self-limited seekers used an average of 29.8 words and self-limited sources 98.8 words. Notice that on the average self-limited seekers used 148 fewer words than did unlimited seekers, while self-limited sources used 268.9 fewer words than did their unrestricted counterparts. It is the difference between the last two figures that accounts for the significant job roles by word usage interaction.

The above pattern of significant main effects for word usage and job roles, plus a significant job roles by word usage interaction, is repeated by many of the dependent measures about to be discussed.

Messages. Unlimited teams used an average of 34.6 messages, while self-limited teams used an average of 16.0 messages. For this dependent measure, the main effect of job roles was not significant, but there was a significant job roles by word usage interaction. Averaged across all unlimited teams, the seekers generated 17.5 messages, and the sources 17.1. Comparable figures for the self-limited teams are 7.8 and 8.2 messages, respectively. Thus, unlimited seekers generated an average of 0.4 messages more than did the sources, while self-limited seekers generated 0.4 messages fewer than did their sources. What this means is that on unlimited teams the seekers generally typed the first and the last message (e.g., "Hello" and "Goodbye"), while on self-limited teams it was the source who frequently had the first and last messages. In the latter case, the seekers, not wanting to waste words, usually quietly waited for the source to begin the instructions, and at the conclusion of the problem turned the system off without a parting message.

Message length. For the measure of message length there was a significant main effect for word usage, a significant main effect for job roles, and a significant job roles by word usage interaction. Unlimited subjects used an average of 18.0 words per message (11.8 for seekers, 24.1 for sources), while the self-limited subjects used an average of only 8.9 words per message (3.6 for the seekers, 14.2 for the sources). Thus, there is essentially the same pattern for message length as there had been for words: On the average, the unlimited subjects used more words per message than did the self-limited, and sources used more than seekers, with the difference between the two job roles being greater in the unlimited condition than in the self-limited.

Notice that none of the figures above can be obtained by dividing the previously given values for words by the reported number of messages. As an illustrative example of why this is the case, let us assume that we have only two subjects, A and B. Let us further assume that A used only one message and one word, and that B used two messages and four words. As a result, A's average message length is one word, and B's is two words. Therefore, averaged across the two subjects, they had a mean of 1.5 words per message. However, if you were to look solely at their mean number of words (2.5) and their mean number of messages (1.5), you would have arrived at a different, incorrect value for their mean number of words per message (1.67 rather than 1.50).

Communication rate. Once again, as was the case with words and message length, there was for communication rate a significant main effect for word usage, a significant main effect for job roles, and a

significant job roles by word usage interaction. Unlimited teams transmitted an average of 20.1 words per minute of problem solution time (6.5 words per minute for the seeker, 13.6 for the source), while self-limited teams transmitted an average of 6.5 words per minute (1.4 for the seeker, 5.1 for the source).

Requests for information. The significant main effects and interactions associated with the different types of information requests can best be explained with the help of Table 7. This table shows the mean number of each type of request generated by seekers and sources in the unlimited and self-limited conditions. The corresponding percentages for each mean were computed by taking that mean value, multiplying by 100, and then dividing by the sum of all of the means in that column. Notice that these percentages do not represent the mean that would have been obtained by averaging the percentages of each individual subject; it was not feasible to obtain a percentage breakdown for each subject because some subjects did not generate any information requests.

All together, the 96 subjects in this study generated a total of 823 requests for information. Of these, only seven were classified as belonging to the category of "other types." The remaining four categories of information requests therefore account for over 99% of all requests.

For each of these four categories the main effect of job roles is highly significant ($p < 0.001$ in all cases). The nature of these differences can easily be seen in Table 7. Further, the differences attributable to the word usage factor can also be seen in Table 7. For all categories, the self-limited teams had fewer information requests than did the unlimited. However, the decrease in the number of requests for clarification failed even to approach significance ($p = 0.34$), although the decreases for the other three were significant. For these latter three there were also significant job roles by word usage interactions, the natures of which can be seen in Table 7.

Discussion. Some of the findings reported in this section are not new. Ford et al. (1979) used a word usage factor very similar to the one I used. I, like they, found that self-limited subjects, when compared with unlimited, use fewer words and fewer messages, average fewer words per message, and have a slower communication rate. This is exactly what I had expected to find. What I had not expected, and what I consider to be one of the most important findings of my study, is that the unlimited subjects took over 33% more time than the self-limited to solve the identical problem. (Ford et al., did not find a significant difference for time to solution. This could be because their experiment used fewer subjects and had less statistical power than mine; also, only half rather than all of their subjects were in a teletypewriter mode.)

Table 7

The Mean Number of Times Subjects in the Two Job Roles and Two Word Usage Conditions Used Each of the Five Types of Information Requests

Type of Information Request	Unlimited		Self-limited	
	Seekers	Sources	Seekers	Sources
Requests for New Information	8.79 (51.7%)	1.00 (13.6%)	3.00 (38.3%)	0.79 (37.3%)
Requests for Clarification of Previously Given Information	2.67 (15.7%)	0.42 (5.7%)	1.83 (23.4%)	0.50 (23.6%)
Requests to Confirm Own Understanding or Status	4.96 (29.2%)	0.08 (1.1%)	2.96 (37.8%)	0.00 (0.0%)
Requests to Confirm Partner's Understanding or Status	0.54 (3.2%)	5.63 (76.7%)	0.04 (0.5%)	0.79 (37.3%)
Other Types	0.04 (0.2%)	0.21 (2.9%)	0.00 (0.0%)	0.04 (1.9%)
TOTAL	17.00 (100%)	7.34 (100%)	7.83 (100%)	2.12 (100%)

NOTE: Each column contains the means of 24 subjects. The values inside the parentheses indicate what percentage of the total number of information requests is accounted for by that particular request.

A clue as to why my self-limited teams were faster may lie in one of the conclusions from the study by Weeks et al. (1974): subjects communicating in a teletypewriter mode tend to transmit messages as they compose them, rather than compose and edit the messages prior to transmission. However, Weeks et al. were talking only about unlimited subjects, and their conclusions do not necessarily apply to subjects who must limit their word usage. This is because, unlike unlimited subjects, self-limited subjects must pay for every word they use, and therefore cannot afford the luxury of "thinking out loud" (i.e. over the teletypewriter). Instead, they tended to compose and edit their messages in advance. The result is that their messages were generally shorter, while apparently conveying as much information as the longer, rambling messages sent by unlimited subjects. Therefore, I believe that the self-limited teams were faster than the unlimited simply because it apparently takes longer to compose a message on a teletypewriter than it does to compose on notepaper or in one's head a concise message and then transmit that.

The above statement might not make intuitive sense to people who find it faster to use a typewriter to compose and edit manuscripts. Keep in mind, though, that a teletypewriter is not nearly as convenient to use as a standard typewriter, in that the user is not permitted to erase anything once it is written, and cannot even return to a previous line to strike something over. This makes composing messages on a teletypewriter a very tedious task.

Let us now move on to examine the requests for information. First, it is important to note that there was <u>not</u> a significant difference in the number of requests for clarification, suggesting that the subjects found unlimited and self-limited dialogue to be equally clear. (In this context, it should be pointed out that the sex of the source had a greater effect than the word usage factor on the number of requests for information.) Second, the self-limited subjects generated far fewer of the four remaining types of requests than did the unlimited subjects. However, as can be seen from the percentages shown in Table 7, some types decreased more dramatically than others.

The greatest decreases are for the seekers' requests for new information and the sources' requests to confirm the partner's understanding or status. Apparently requests of these types are a normal feature of unrestricted dialogue that are, for the most part, unnecessary for the exchange of factual information.

<u>Sex of Seeker by Sex of Source by Anonymity</u>

The multivariate F for the sex of seeker by sex of source by anonymity interaction was significant, but barely so ($\underline{p} < 0.049$). The only significant univeriate F test was on the questionnaire ratings of sources' brevity. In describing the nature of this three-way interaction it is more convenient, and just as accurate, to think of it as a

simple two-way interaction: the anonymity factor by the factor of like- versus mixed-sex teams.

In the like-sex conditions (i.e., all-male and all-female) the introduced sources received a lower average rating than did the anonymous sources (4.2 versus 5.2 on the seven-point rating scale). On the other hand, the sources on the introduced mixed-sex teams received higher ratings than their anonymous counterparts (5.2 versus 4.7).

I am unable to offer what I would regard as a satisfactory explanation for this interaction. None of the other 14 univariate analyses reached significance for this source of variation, so it is impossible to determine a firm basis that would account for the sources being rated differently in the different conditions. Because the multivariate F barely reached significance, and because only one of the univariate F's reached significance (again just barely), I am inclined to believe that this was simply a chance effect.

Job Roles

There are two significant main effects for job roles in addition to the seven discussed in the Word Usage section. These are for the questionnaire ratings of the seekers' clarity and brevity.

Seekers' clarity. In rating the seekers' clarity on the seven-point rating scale, seekers gave themselves an average rating of 5.8, while the sources gave them an average rating of 6.5 ($p < 0.001$). Unlimited sources gave slightly higher ratings to seekers than did the self-limited sources (6.7 versus 6.2), causing the job roles by word usage interaction to also reach significance ($p < 0.05$).

Seekers' brevity. In ratings of the seekers' brevity, seekers gave themselves an average rating of 5.9, while receiving from the sources an average rating of 6.4 ($p < 0.01$).

Discussion. In general, seekers did not rate their clarity or brevity as highly as did the sources. This could simply be due to modesty on the part of the seekers. It is interesting to note that there were no significant job roles main effects for the ratings of the sources' performances, perhaps because the sources used many more words than the seekers, thereby providing a firmer basis for judgment.

In addition, on the average the sources apparently felt that the seekers were slightly clearer in the unlimited than in the self-limited condition. Recall, however, that this is not reflected by a significant difference in the number of requests for clarification.

Intercorrelations of Dependent Measures

The within-cell intercorrelation matrix for the 15 dependent measures is shown in Table 8. Although the values in Table 8 can be

Table 8

The Within-cell Intercorrelations of the 15 Dependent Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) Time to Solution	-----														
(2) Words	.460	-----													
(3) Messages	.406	.153	-----												
(4) Message Length	-.073	.441	-.721	-----											
(5) Communication Rate	-.307	.597	-.123	.417	-----										
(6) New Information	.281	-.008	.785	-.637	-.309	-----									
(7) Clarification	.412	.378	.467	-.066	.137	.058	-----								
(8) Own Understanding	.259	-.066	.661	-.589	-.168	.317	.252	-----							
(9) Partner's Understanding	.009	.341	.418	-.216	.373	.279	.105	.186	-----						
(10) Other Types	.139	-.089	.492	-.448	-.195	.545	-.095	.325	.372	-----					
(11) Seekers' Clarity	-.118	-.267	-.172	-.032	-.167	.212	-.379	-.233	-.317	-.066	-----				
(12) Seekers' Brevity	-.397	-.461	-.303	.018	-.119	-.096	-.485	-.105	-.083	.104	.372	-----			
(13) Sources' Clarity	-.366	-.286	-.062	-.120	.028	.070	-.456	-.074	.114	.364	.324	.561	-----		
(14) Sources' Brevity	-.224	-.080	-.156	.056	.035	.036	-.289	-.145	.114	.153	.106	.295	.587	-----	
(15) Team Performance	-.224	.069	-.199	.231	.127	.108	-.317	-.394	.116	-.119	.267	.062	.149	.356	-----

NOTE: The above $\underline{r}$'s have 31 degrees of freedom. The absolute value of the $\underline{r}$'s must exceed 0.344 for $\underline{p} \leq 0.05$, and 0.443 for $\underline{p} \leq 0.01$.

interpreted as though they are standard Pearson product-moment correlation coefficients (r), they were computed in a different manner. This is because there were 16 different factor-level combinations, or cells, in this experiment (see Table 1), and each of these combinations of experimental factors had its own influence on the dependent measures. These influences are manifested by the differences among the cell means. (These differences among cell means are, of course, what is being measured by analysis of variance.) Clearly, the differences among the cell means can bias the computed intercorrelations of the dependent measures.

The computing of correlation coefficients has as one of the steps the conversion of all raw scores to standard scores (or z-scores). In order to eliminate the influences of the experimental factors on the dependent measures, I did not compute the r's in Table 8 using standard scores that had been based on the grand mean of each distribution; rather, I based the standard scores on their corresponding cell means.

This procedure compensated for the differences in the cell means, but at a cost of statistical power, as measured by degrees of freedom. Typically, r's have for their degrees of freedom the total number of pairs of observations, minus two degrees of freedom, one of which is for the mean. In this case, 16 different means were used (one for each cell), and each one cost a degree of freedom. The total number of degrees of freedom for the r's in Table 8 therefore is 31, which equals the number of teams (48), minus the number of cells (16), minus one.

Given that the r's in Table 8 have 31 degrees of freedom, their absolute values must therefore exceed 0.344 to be significant beyond the 0.05 level, and 0.443 to be significant beyond the 0.01 level. Out of the 105 correlation coefficients shown in Table 8, 31 are significant beyond the 0.05 level, 16 of which are significant beyond the 0.01 level. This very large proportion of significant correlations confirms the importance and value of using a MANOVA. However, the information contained in Table 8 does more than simply justify the use of MANOVA. It provides, for the first time in a Hopkins telecommunications experiment, an indication of which of the dependent measures are correlated with each other and which are independent. There were some surprises.

As seems intuitively reasonable, time to solution was positively and significantly correlated with both number of words and number of messages. However, what was unexpected is that the number of words and the number of messages are almost entirely independent of each other. Message length (simply the number of words divided by the number of messages) is independent of time to solution, but, as would be expected, it is significantly positively correlated with number of words, and strongly negatively correlated with number of messages. Communication rate (number of words divided by time to solution) is positively correlated with both number of words and message length.

So what does all this tell us? It tells us that teams, when they take longer, tend to use more words and more messages than teams that did not take as long. However, just because one team may have used more words than another, it does not necessarily mean that it also used more messages. This is because the average message length tends to increase, and not the number of messages. As a corollary to this, when a team does use a greater number of messages, the messages tend to be shorter. Further, the teams that tended to use more words, and therefore longer messages, also tended to have a faster communication rate.

Turning to the requests for information, the teams with the higher number of requests for clarification also tended to take longer and use more words and messages. None of the other types of information requests were significantly correlated with time to solution and number of words, although all of them were significantly correlated with the number of messages. This is hardly surprising in light of the fact that each of these requests frequently constituted a message.

There are some interesting points to be made with regard to the questionnaire data. Of the 15 correlations of the five questions with time, words, and messages, 14 are negative. This means that the more time, words, and messages a team used, the lower their questionnaire ratings tended to be. Further, the more requests for clarification and requests to confirm own understanding there were, the lower the ratings tended to be, especially with regard to the requests for clarification. It is interesting to note, however, that the number of the other types of requests, as well as message length and communication rate, are uncorrelated with the subjects' perceptions of performance.

Finally, the response to each question was significantly and positively correlated with the responses given to the adjacent questions on the questionnaire. This suggests that the answer given to a question may have been affected by the answer given to the previous one, and that I should have used different forms of the questionnaire, with the questions in a different order on each form.

Mistakes Made

The mistakes I made with this experiment were, for the most part, errors of omission. There were some potentially useful and easily obtainable dependent measures I failed to examine simply because their use had not occurred to me in advance.

As has been previously discussed, useful information might have been obtained by keeping track of the number of times the seeker improperly assembled and then disassembled the model. Also, a more accurate measure of communication rate could have been obtained had I kept track of the subjects' actual communication time.

Finally, I noticed that some seekers turned off the system power immediately after assembling the model, while others took as much as ten minutes examining the model and asking questions before deciding that they were done. I have no records to indicate whether the seekers took longer in some conditions than in others to examine their finished models. Further, not only might there have been differences among the conditions in how much time the seekers took to check out their completed models, but there might also be differences in the number of words, messages, and requests for information. This might have proven to be extremely valuable data. For example, based solely on the data I did collect, it is possible that the unlimited and the self-limited team did not differ significantly while assembling the model; however, as soon as the model was assembled, the unlimited teams took the time to make sure it was correct, while the self-limited teams, not wanting to waste words, stopped working. Although I am reasonably certain that this is not what happened, I cannot prove it.

Thus, I feel that the mistakes I made in conducting this experiment were, for the most part, my failures to recognize, except in retrospect, the need for the above-mentioned dependent measures.

Difficulties Encountered

I encountered only one major difficulty while conducting this experiment; I had an extremely difficult time obtaining subjects, especially women. There were 2016 full-time undergraduate students at Johns Hopkins the semester I conducted my experiment, 1419 men and 597 women. Finding 48 men who were willing to participate in the experiment was a nuisance; finding 48 women proved to be a real challenge.

Although I had no trouble finding subjects when I first began the experiment, toward the end I was forced, among other things, to interrupt classes being taught by fellow graduate students (with their permission, of course) to ask for volunteers. I continually put up my advertising posters wherever on the campus I thought they might do some good, even to include the shower rooms in the women's dormitories. In short, I had an extremely difficult time making sure that I had enough subjects, especially women. I am convinced that it would have been impossible for me to conduct an experiment at Hopkins that needed many more than the 48 women mine used.

I also encountered a couple of minor problems. Early in the experiment I found that the subjects had a tendency to forget either the time or the room number I had assigned to them. What I wound up doing was, after scheduling a subject by phone, putting a note in his or her campus mailbox telling the subject when and where to report for the experiment. The notes also included my name and campus telephone number. This cut down on absenteeism tremendously. Also, those who found that they could not come at the assigned time often called me well enough in advance for me to conveniently reschedule them.

The second minor problem involved the anonymous teams. In talking to the subjects during their pre-session briefings, I had to remember to refer to their partners as "your partner" and not by pronouns such as he or she. It took a few sessions before I could do this smoothly.

So, to sum up, I feel that the only major difficulty involved with conducting this experiment was due to the small size of the available subject pool, causing me to expend a tremendous amount of effort obtaining volunteers. The minor difficulties involved getting the subjects to remember to show up, and getting me to remember not to accidentally give extra information to the anonymous subjects.

Conclusion

In this section I shall first quickly summarize the important findings of my study, and then discuss where these findings lead with regard to potential areas for future research.

Summary

One of the purposes of this experiment was to determine whether or not there is a mismatch in the way men and women communicate to exchange factual information and solve problems. No evidence of such a mismatch was found, suggesting that the difference in performance between like- and mixed-sex teams in previous studies of this type was due to the confounding influences of non-verbal communication. From an applied standpoint, this is strong evidence that natural language computers will not have to adjust themselves to the sex of their user.

Further, whether the subjects were previously introduced or not apparently had no effect on their being able to work together as a team.

There were two major findings in this study in which condition-attributable differences were demonstrated. The first is that the sex of the person receiving instructions could not be shown to have any effect on the team's performance, although performance was affected by the sex of the person giving instructions. The females proved to be superior on all measures where there was a significant difference. I believe that, for the most part, this can be attributed simply to the females being more thorough and careful than the males.

The second major finding is that when the subjects in this study had an incentive to cut down on their word usage, they not only used fewer words, fewer messages, and averaged shorter message lengths than the unlimited subjects, but they also solved their problems faster. I believe that the major reason for this finding is that the self-limited subjects, in general, composed and carefully edited their

messages prior to transmission, while the unlimited subjects tended to compose their messages "on-line."

Future Research

Although this experiment answered the applied questions it had been intended to answer, it has raised many new theoretical questions that can be answered only by further experimentation.

With regard to the findings on sex differences: How robust are they? In other words, are females generally superior at giving instructions? Or are there certain kinds of problems on which males would be superior? If so, what kinds of problems? Further, are there certain kinds of problems that would, unlike the problem I used, result in a communications mismatch between the sexes? If so, again what kinds of problems? There are many many more questions in addition to these. The point is that my experiment was anything but the definitive theoretical work on sex differences in interactive verbal telecommunication, but at least it was a start.

The findings with regard to self-limited versus unlimited word usage also suggest the need for further research, especially because of the possibility that the self-limited teams were faster than the unlimited because their method of communication was less natural. Klemmer (1973), in comparing the merits of standard versus video-supplemented telephones, warned against assuming that the more natural the method of telecommunication, the better the method is. Until now, to the best of my knowledge, no laboratory experiment has actually demonstrated a significant improvement in efficiency attributable to a decrease in the convenience of communication within the same mode (e.g., teletypewriter). However, I believe that my self-limited teams were faster than the unlimited at least in part because of their inability to communicate in a natural manner. Further, the mixed-sex teams were as efficient as the like-sex almost certainly because of the absence of the non-verbal capacity that is available in the more natural communication modes. I therefore feel that some benefit may be derived from future research into how "natural" a communication system should be in each situation in order to exchange information in the most efficient manner. My findings suggest that by following this course, we might, as Licklider et al. (1968) predicted, "be able to communicate more effectively through a machine than face to face."

References

Astin, H. S. Sex differences in mathematical and scientific precocity. In Stanley, J. C., Keating, D. P. & Fox, L. H. (eds.). Mathematical talent: Discovery, description and development. Baltimore: The Johns Hopkins University Press, 1974.

Barclay, A. M. The effect of female aggressiveness on aggressive and sexual fantasies. Journal of Projective Techniques & Personality Assessment, 1970, 34, 19-26.

Barfield, A. Biological influences on sex differences in behavior. In Teitelbaum, M. S. (ed.). Sex differences: Social and biological perspectives. Garden City, New York: Anchor Press, 1976.

Bedell, J. & Sistrunk, F. Power, opportunity costs, and sex in a mixed-motive game. Journal of Personality and Social Psychology, 1973, 25, 219-226.

Bieri, J., Bradburn, W. & Galinsky, M. Sex differences in perceptual behavior. Journal of Personality, 1958, 26, 1-12.

Blackwell, A. M. A comparative investigation into the factors involved in mathematical ability of boys and girls. British Journal of Educational Psychology, 1940, 10, 143-153.

Carey, G. L. Sex differences in problem-solving performance as a function of attitude differences. Journal of Abnormal and Social Psychology, 1958, 56, 256-260.

Chapanis, A. Prelude to 2001: Explorations in human communication. American Psychologist, 1971, 26, 949-961.

Chapanis, A. The communication of factual information through various channels. Information Storage and Retrieval, 1973, 9, 215-231.

Chapanis, A., Ochsman, R. B., Parrish, R. N. & Weeks, G. D. Studies in interactive communication: I. The effects of four communication modes on the behavior of teams during cooperative problem-solving. Human Factors, 1972, 14, 487-509.

Chapanis, A. & Overby, C. M. Studies in interactive communication: III. Effects of similar and dissimilar communication channels and two interchange options on team problem solving. Perceptual and Motor Skills, Monograph Supplement 2-V38, 1974, 38, 343-374.

Chapanis, A., Parrish, R. N., Ochsman, R. B. & Weeks, G. D. Studies in interactive communication: II. The effects of four communication modes on the linguistic performance of teams during cooperative problem solving. Human Factors, 1977, 19, 101-126.

Churchill, T. O. The life of Lord Viscount Nelson. London, England: J. and W. MacGavin, 1808.

Clarke, A. C. 2001: A space odyssey. New York: New American Library, 1968.

Crandall, V. C. Sex differences in expectancy of intellectual and academic reinforcement. In Smith, C. P. (ed.). Achievement related motives in children. New York: Russell Sage Foundation, 1969.

Deaux, K. Ahhh, she was just lucky. Psychology Today, 1976, 10 (6), 70-75.

Doob, A. N. & Gross, A. E. Status of frustrator as an inhibitor of horn-honking responses. Journal of Social Psychology, 1968, 76, 213-218.

Droege, R. C. Sex differences in aptitude maturation during high school. Journal of Counseling Psychology, 1967, 14, 407-411.

Ehrhardt, A. A. & Baker, S. W. Hormonal aberrations and their implications for the understanding of normal sex differentiation. Paper presented at the meetings of the Society for Research in Child Development, Philadelphia, 1973.

Erlich, A. C. & LeBold, W. K. Factors influencing the science career plans of high school students. Measurement and Research Center, Purdue University, West Lafayette, Indiana, June 1975.

Feather, N. T. Change in confidence following success or failure as a predictor of subsequent performance. Journal of Personality and Social Psychology, 1968, 9, 38-46.

Feshbach, S. Aggression. In Mussen, P. H. (ed.). Carmichael's manual of child psychology. New York: Wiley, 1970.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr, D. B., Goldberg, I. & Neyman, C. A. Counselor's technical manual for interpreting test scores (Project Talent), Palo Alto, California, 1961.

Ford, W. R., Chapanis, A. & Weeks, G. D. Self-limited and unlimited word usage during problem solving in two telecommunication modes. Journal of Psycholinguistic Research, 1979. In press.

Fox, L. H., Pasternak, S. R. & Peiser, H. L. Career related interests of adolescent boys and girls. In Keating, D. P. (ed.). Intellectual talent: Research and development. Baltimore, Maryland: The Johns Hopkins University Press, 1976.

Gates, A. I. Sex differences in reading ability. Elementary School Journal, 1961, 61, 431-434.

Harms, L. S. Human communication: The new fundamentals. New York: Harper & Row, 1974.

Hoecker, D. G. Problem-solving in five communication modes as a function of verbal and spatial abilities. Doctoral dissertation, The Johns Hopkins University, 1979.

Holland, J. L. The Vocational Preference Inventory. Palo Alto, California: Consulting Psychologists Press, 1965.

Keating, D. P. The study of mathematically precocious youth. In Stanley, J. C., Keating, D. P. & Fox, L. H. (eds.). Mathematical talent: Discovery, description and development. Baltimore, Maryland: The Johns Hopkins University Press, 1974.

Kelly, M. J. & Chapanis, A. Limited vocabulary natural language dialogue. International Journal of Man-Machine Studies, 1977, 9, 497-501.

Klemmer, E. T. Interpersonal communication systems: Relevance, credibility, impact. Division 21, Presidential Address, American Psychological Association, Montreal, Canada, August, 1973.

Licklider, J. C. R., Taylor, R. W. & Herbert, E. The computer as a communication device. Science and Technology, 1968, 76, 21.

Maccoby, E. E. & Jacklin, C. N. The Psychology of Sex Differences. Stanford, California: Stanford University Press, 1974.

McIntyre, A. Sex differences in children's aggression. Proceedings of the 80th Annual Convention of the A.P.A., 1972, 7, 93-94.

McNeal, S. P., McClintock, C. G., & Nuttin, J. M. Effects of sex role in a two person mixed-motive game. Journal of Personality and Social Psychology, 1972, 24, 372-380.

Michaelis, P. R., Chapanis, A., Weeks, G. D. & Kelly, M. J. Word usage in interactive dialog with restricted and unrestricted vocabularies. IEEE Transactions on Professional Communication, 1977, 20, 214-221.

Monday, L. A., Hout, D. P. & Lutz, S. W. College student profiles. American College Testing Program. Iowa City: ACT Publications, 1966-67.

Money, J. & Ehrhardt, A. A. Man and woman, boy and girl. Baltimore, Maryland: The Johns Hopkins University Press, 1972.

Ochsman, R. B. & Chapanis, A. The effects of 10 communication modes on the behavior of teams during co-operative problem-solving. International Journal of Man-Machine Studies, 1974, 6, 579-619.

Parrish, R. N. Interactive communication in team problem-solving as a function of two educational levels and two communication modes. Doctoral dissertation, The Johns Hopkins University, 1973. Dissertation Abstracts International, 1974, 34, 5721B. (University Microfilms, No. 74-10, 440.)

Parrish, R. N. The performance of college freshmen and seniors during cooperative problem solving. International Journal of Man-Machine Studies, 1978, 10, 553-568.

Rapoport, A. & Chammah, A. M. Sex differences in factors contributing to the level of cooperation in the Prisoner Dilemma Game. Journal of Personality and Social Psychology, 1965, 2, 831-838.

Rosenberg, B. G. & Sutton-Smith, B. The relationship of ordinal position and sibling sex status to cognitive abilities. Psychonomic Science, 1964, 1, 81-82.

Rosenberg, B. G. & Sutton-Smith, B. Sibling association, family size, and cognitive abilities. Journal of Genetic Psychology, 1966, 109, 271-279.

Rosenberg, B. G. & Sutton-Smith, B. Sibling age spacing effects upon cognition. Developmental Psychology, 1969, 1, 661-668.

Rosenthal, R. Experimenter effects in behavioral research. New York: Irvington Publishers, Inc., 1976.

Rychlak, J. F. & Lerner, J. J. An expectancy interpretation of manifest anxiety. Journal of Personality and Social Psychology, 1965, 2, 667-684.

Schwartz, S. H. & Clausen, G. T. Responsibility, norms and helping in an emergency. Journal of Personality and Social Psychology, 1970, 16, 299-310.

Shomer, R. W., Davis, A. H. & Kelley, H. H. Threats and the development of coordination: Further studies in the Deutsch and Krauss trucking game. Journal of Personality and Social Psychology, 1966, 4, 119-126.

Stewart, V. Social influences on sex differences in behavior. In Teitelbaum, M. S. (ed.). Sex differences: Social and biological perspectives. Garden City, New York: Anchor Press, 1976.

Svensson, A. Relative achievement. School performance in relation to intelligence, sex and home environment. Stockholm: Almquist & Wiksell, 1971.

Tedeschi, J. T., Lesnick, S. & Gahagan, J. P. Feedback and "washout" effects in the Prisoner's Dilemma game. Journal of Personality and Social Psychology, 1968, 10, 31-34.

Terman, L. M. Mental and physical traits of a thousand gifted children. Genetic studies of genius, vol. 1. Stanford, California: Stanford University Press, 1925.

Very, P. S. Differential factor structures in mathematical abilities. Genetic Psychology Monographs, 1967, 75, 169-207.

Walberg, H. J. Physics, femininity and creativity. Developmental Psychology, 1969, 1, 47-54.

Warner, O. Great Sea Battles. London, England: Hamlyn Publishing Group, Ltd., 1968.

Weeks, G. D., Kelly, J. J. & Chapanis, A. Studies in interactive communication: V. Cooperative problem solving by skilled and unskilled typists in a teletypewriter mode. Journal of Applied Psychology, 1974, 59, 665-674.

Youssef, Z. I. The role of race, sex, hostility, and verbal stimulus in inflicting punishment. Psychonomic Science, 1968, 12, 285-286.

TECHNICAL REPORTS DISTRIBUTION LIST

Code 455

CDR Paul R. Chatelier
Office of the Deputy Under Secretary
of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Director
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217 (5cys)

Director
Analysis and Support Division
Code 230
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Operations Research Programs
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Information Systems Program
Code 437
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Branch Office
ATTN: Dr. J. Lester
Building 114, Section D
666 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. Davis
536 South Clark Street
Chicago, IL 60605

Commanding Officer
ONR Branch Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, CA 96503

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6cys)

Dr. Bruce Wald
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D.C. 20375

Dr. Robert G. Smith
Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Washington, D.C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-00T
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. Warren Lewis
Human Engineering Branch
Code 8231
Naval Ocean Systems Center
San Diego, CA 92152

CAPT Richard L. Martin, USN
Commanding Officer
USS Francis Marion (LPA-249)
FPO New York 09501

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 08T24
Washington, D.C. 20360

Commander
Naval Air Systems Command
Human Factors Programs
NAVAIR 340F
Washington, D.C. 20361

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, D.C. 20362

Naval Sea Systems Command
Personnel & Training Analyses Office
NAVSEA 074C1
Washington, D.C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, D.C. 20360

Mr. Milon Essoglou
Naval Facilities Engineering Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

Scientific Advisor to DCNO (MPT)
OP 01T (Dr. Marshall)
Washington, D.C. 20370

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Navy Personnel Research and Development Center
Manned Systems Design, Code 311
San Diego, CA 92152

Navy Personnel Research and Development Center
Code 305
San Diego, CA 92152

Navy Personnel Research and Development Center
Management Support Department
Code 210
San Diego, CA 92151

CDR P. M. Curran
Code 604
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, MD 21402

Human Factor Engineering Branch
Naval Ship Research and Development
Center, Annapolis Division
Annapolis, MD 21402

Mr. Merlin Malehorn
Office of the Chief of Naval
Operations (OP 102X)
Washington, D.C. 20350

Mr. J. Barber
HQS, Department of the Army
DAPE-MBR
Washington, D.C. 20310

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and
Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Edgar M. Johnson
Organizations and Systems Research
Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

U.S. Army Aeromedical Research Lab
ATTN: CPT Gerald P. Krueger
Ft. Rucker, AL 36362

ARI Field Unit-USAREUR
ATTN: Library
C/O ODCSPER
HQ USAREUR & 7th Army
APO New York 09403

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Air University Library
Maxwell Air Force Base, AL 36112

Dr. Gordon Eckstrand
AFHRL/ASM
Wright-Patterson AFB, OH 45433

North East London Polytechnic
The Charles Myers Library
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

Professor Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology
Establishment
Teddington, Middlesex TW11 OLN
ENGLAND

Director, Human Factors Wing
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 cys)

Dr. Stephen J. Andriole
Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. Stanley Deutsch
Office of Life Sciences
National Aeronautics and Space Administration
600 Independence Avenue
Washington, D.C. 20546

Dr. Robert R. Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, CA 93017

Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Dr. H. McIlvaine Parsons
Institute for Behavioral Research
2429 Linden Lane
Silver Spring, MD 20910

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert Williges
Human Factors Laboratory
Virginia Polytechnical Institute and State University
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, NW
Washington, D.C. 20036

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, NW
Washington, D.C. 20036 (3cys)

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics Company
St. Louis Division
Box 516
St. Louis, MO 63166

ing that the difference in performance
teams in previous studies of this type was
uences of non-verbal communication. From
is strong evidence that natural language
adjust themselves to the sex of their user.

e subjects were previously introduced or not
their being able to work together as a

r findings in this study in which condition-
re demonstrated. The first is that the sex
tructions could not be shown to have any
mance, although performance was affected by
g instructions. The females proved to be
ere there was a significant difference. I
part, this can be attributed simply to the
and careful than the males.

nding is that when the subjects in this
ut down on their word usage, they not only
sages, and averaged shorter message lengths
, but they also solved their problems
major reason for this finding is that the
eneral, composed and carefully edited their